



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 117804

To: Dave Nguyen
Location: rem/2d31/2c18
Art Unit: 1632
Monday, March 29, 2004

Case Serial Number: 09/869508

From: Beverly Shears
Location: Remsen Bldg.
RM 1A54
Phone: 571-272-2528

beverly.shears@uspto.gov

Search Notes

SEARCH REQUEST FORM

Requestor's
Name: _____

Serial
Number: _____

Date: _____ Phone: _____ Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: 03-29-04
Searcher: Beverly C 2528
Terminal time: 20
Elapsed time: _____
CPU time: _____
Total time: 23
Number of Searches: _____
Number of Databases: 1

Search Site

____ STIC
____ CM-1
____ Pre-S
Type of Search
____ N.A. Sequence
____ A.A. Sequence
____ Structure
____ Bibliographic

Vendors

____ IG
____ STN
____ Dialog
____ APS
____ Geninfo
____ SDC
____ DARC/Questel
☒ Other CGN

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2004, 12:40:25 ; Search time 6740 Seconds
(without alignments)
10919.358 Million cell updates/sec

Title: US-09-869-508-1_COPY_453_2150

Perfect score: 1698

Sequence: 1 agaacattatggtggccaa.....caccocggaggagccgcag 1698

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_ste.*
28: em_un.*
29: em_vi.*
30: em_hgt_hum.*
31: em_hgt_inv.*
32: em_hgt_other.*
33: em_hgt_mus.*
34: em_hgt_pln.*
35: em_hgt_rod.*
36: em_hgt_man.*
37: em_hgt_vrt.*
38: em_sy.*
39: em_hgtgo_hum.*
40: em_hgtgo_mus.*
41: em_hgtgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1698	100.0	2550	9	HSYB1	X96666 H.sapiens Y
2	1672.4	98.5	161835	9	AC098484	AC098484 Homo sapi
3	456.6	26.9	659	9	HUMYZ92E04	AF086102 Homo sapi
C 4	186.8	11.0	162454	10	EX470216	EX470216 Mouse DNA
C 5	186.8	11.0	194020	2	AC101988	AC101988 Mus muscu
C 6	186.2	11.0	225308	2	AC098918	AC098918 Rattus no
7	111	6.5	178	6	AX897942	AX897942 Sequence
8	111	6.5	178	6	BD033475	BD033475 Sequence
9	107.4	6.3	1548	6	AX014868	AX014868 Sequence
10	107.4	6.3	1548	6	BD037113	BD037113 Human nuc
11	101.4	6.0	144902	2	AC016114	AC016114 Homo sapi
12	101.4	6.0	162144	9	CNS06C85	AL391261 Human chr
C 13	101.4	6.0	184656	9	CNS05TDB	AL356021 Human chr
C 14	96.6	5.7	189367	9	AL162733	AL162733 Human DNA
C 15	95	5.6	99886	9	AL135841	AL135841 Human DNA
16	89.4	5.3	395	6	AX321258	AX321258 Sequence
17	89	5.2	1481	6	AR117695	AR117695 Sequence
18	89	5.2	1481	6	AX409453	AX409453 Sequence
19	89	5.2	1481	6	AX498410	AX498410 Sequence
20	89	5.2	1481	6	AX840171	AX840171 Sequence
21	89	5.2	1481	9	HUMYB1A	J03827 Y-box bindi
22	87.4	5.1	2168	9	HUMPSDBPB	L37516 Homo sapien
23	86.6	5.1	162996	2	AC051655	AC051655 Homo sapi
24	85.2	5.0	410	6	AX321869	AX321869 Sequence
25	84	4.9	1521	9	BC038384	BC038384 Homo sapi
26	83	4.9	3073	6	AR083634	AR083634 Sequence
27	83	4.9	3073	6	122491	122491 Sequence 31
28	82	4.8	917	9	BC000064	BC000064 Homo sapi
29	81.4	4.8	1468	9	HUMENSEP	M83234 Human nucle
30	81.4	4.8	1474	9	HUMENABP	M85234 Human nucle
C 31	80.8	4.8	16805	2	AC147392	AC147392 Otolenur
32	80.6	4.7	142184	2	AC037428	AC037428 Homo sapi
33	80.6	4.7	143604	9	AC093610	AC093610 Homo sapi
34	80.6	4.7	310021	2	AC092883	AC092883 Homo sapi
C 35	79	4.7	77948	2	AC018982	AC018982 Homo sapi
36	78.8	4.6	381	6	BD275806	BD275806 COMPOUNDS
37	78.8	4.6	381	6	AR220591	AR220591 Sequence
38	78.8	4.6	381	6	AR255585	AR255585 Sequence
39	78.8	4.6	381	6	AR281155	AR281155 Sequence
40	78.8	4.6	381	6	AX365822	AX365822 Sequence
C 41	77.8	4.6	242756	2	AC093965	AC093965 Rattus no
42	77.4	4.6	545	6	AX321285	AX321285 Sequence
43	77	4.5	1554	9	BC010430	BC010430 Homo sapi
44	76.8	4.5	142565	10	AL592545	AL592545 Mouse DNA
45	76.8	4.5	197658	2	AL646094	AL646094 Mus muscu

ALIGNMENTS

RESULT 1

HSYB1 HSYB1 HSYB1
LOCUS H.sapiens YB-1 gene promoter region.
DEFINITION X96666
ACCESSION X96666
VERSION X96666.1 GI:1403348
KEYWORDS promoter region; Y box binding protein; YB-1 gene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Makino, Y., Ohga, T., Toh, S., Koike, K., Okumura, K., Wada, M.,
Kuwano, M. and Kohno, K.
TITLE Structural and functional analysis of the human Y-box binding

Pred. No. is the number of results predicted by chance to have a

protein (YB-1) gene promoter
Nucleic Acids Res. 24 (10), 1873-1878 (1996)
MEDLINE 96226173
PUBMED 8657568
REFERENCE 2 (bases 1 to 2550)
AUTHORS Yoshinari M.
TITLE Direct Submission
JOURNAL Submitted (15-MAR-1996) Yoshinari M., Department of Biochemistry,
Kyushu University School of Medicine, Maidashi, Fukuoka, 812-82,
JAPAN

FEATURES
source
1..2550
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
gene
1..2411
/gene="YB-1"
promoter
1..1855
/gene="YB-1"
exon
1856..2352
/gene="YB-1"
/number=1
repeat_region
1904..1914
/rpt_type="INVERTED"
protein_bind
1957..1962
/gene="YB-1"
/bound_moiety="SP-1"
protein_bind
2001..2006
/gene="YB-1"
protein_bind
2024..2029
/bound_moiety="SP-1"
protein_bind
2033..2038
/gene="YB-1"
repeat_region
2034..2044
/rpt_type="INVERTED"
protein_bind
2100..2105
/gene="YB-1"
CDS
2187..2411
/gene="YB-1"
/codon_start=1
/protein_id="CAA65446.1"
/db_xref="GI:1403349"
/translation="MSSEAEIQPPAPPAALSAADTKPGTTGSGAGSGPGGLTS
APAGGDKRVIGSDROGWMGPRAAQRNR"

ORIGIN
Query Match 100.0%; Score 1698; DB 9; Length 2550;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1698; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGAACCATATGGTCGGCAAGCTTATATTATTAATTTGGCTCTCACAGAAAAGTT 60
Db 453 AGAACCATATGGTCGGCAAGCTTATATTATTAATTTGGCTCTCACAGAAAAGTT 512
Qy 61 CCTACCCCTTGGCTTAAGCAATTGGTGAATGCGCTACAGTAGTGACAGTAGAATGAA 120
Db 513 CCTACCCCTTGGCTTAAGCAATTGGTGAATGCGCTACAGTAGTGACAGTAGAATGAA 572
Qy 121 AAGACTCTGAATGACATTTAGAGGTTTAGGACTGACTCAAGAGCCCTTTAGGAGGTG 180
Db 573 AAGACTCTGAATGACATTTAGAGGTTTAGGACTGACTCAAGAGCCCTTTAGGAGGTG 632
Qy 181 GAACCTCAGGCCCTAGACGGCATTTGGTAGGGGTAAAGCAATACCTAGCTTACCTTAAC 240
Db 533 GAACCTCAGGCCCTAGACGGCATTTGGTAGGGGTAAAGCAATACCTAGCTTACCTTAAC 692
Qy 241 GCTTTCTCTACTCCATTCCTTGGCCCTCTGCAACCCCAATTCACCTCCGACGCCATTTT 300

693 GCTTTCTCTACTCCATTCCTTGGCCCTCTGCAACCCCAATTCCTCCATCCGACGCCATTTT 752
301 TAAAGAGATGCCCTCTCTACTTATGACTCTAAAATTTGCTCTTCTCCTCCTCTCCCTCAG 360
753 TAAAGAGATGCCCTCTCTACTTATGACTCTAAAATTTGCTCTTCTCCTCCTCTCCCTCAG 812
361 GATATATTTCCAATTAATATACCTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 420
813 GATATATTTCCAATTAATATACCTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 872
421 TCGAGTCTTACTGAATCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
873 TCGAGTCTTACTGAATCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 932
481 CTTGTGCGCGGAAAGTCAAGCTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
933 CTTGTGCGCGGAAAGTCAAGCTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 992
541 GGTTCGAGCTCTGGAGTATTTACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
993 GGTTCGAGCTCTGGAGTATTTACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1052
601 GGCAGAACTGACTCGCGAACTATTCATCCCAAGCGGATAGTAGCGCTTAAAGAAAGAA 660
1053 GGCAGAACTGACTCGCGAACTATTCATCCCAAGCGGATAGTAGCGCTTAAAGAAAGAA 1112
661 CGGAGAGGTTGGTGGGAGGACTTCAGTAACATCAGTGCAGCTCAATTTTATCGTT 720
1113 CGGAGAGGTTGGTGGGAGGACTTCAGTAACATCAGTGCAGCTCAATTTTATCGTT 1172
721 TGTGAAACGTTGATAGTAATCCCTCTATCACGTGGCTGTTGCGAGGATAAAGTGAAGAA 780
1173 TGTGAAACGTTGATAGTAATCCCTCTATCACGTGGCTGTTGCGAGGATAAAGTGAAGAA 1232
781 CAAACAGGCTAGCTTCTCAATAAATGAGTGAATTAATTAATTAATTAATTAATTAATTA 840
1233 CAAACAGGCTAGCTTCTCAATAAATGAGTGAATTAATTAATTAATTAATTAATTAATTA 1292
841 GAAAGAGATGTAATCTCGAAGAGGAGACACATTTTAAATATATATATATATATATATAT 900
1293 GAAAGAGATGTAATCTCGAAGAGGAGACACATTTTAAATATATATATATATATATATAT 1352
901 CGGATCAGAGGCTAGCTCCCTATCGAGACACACCTCCCTTAAACATGCTGAAACCCGGGC 960
1353 CGGATCAGAGGCTAGCTCCCTATCGAGACACACCTCCCTTAAACATGCTGAAACCCGGGC 1412
961 TGCCATAGCTGCTGGTGGTCCCTCCAGGTGACTGCTCCGACAAAGGAGTACGCTCTTCAA 1020
1413 TGCCATAGCTGCTGGTGGTCCCTCCAGGTGACTGCTCCGACAAAGGAGTACGCTCTTCAA 1472
1021 ACGCATACGTTTAAAGCAATTCAGAAACCTCGGCTGTGCGCGGACTACACGGCCATTA 1080
1473 ACGCATACGTTTAAAGCAATTCAGAAACCTCGGCTGTGCGCGGACTACACGGCCATTA 1532
1081 AAGAAAGACACATCTATGCGCGGCTAAGTCTCAGATCAGAGGACCGTATTTGGAG 1140
1533 AAGAAAGACACATCTATGCGCGGCTAAGTCTCAGATCAGAGGACCGTATTTGGAG 1592
1141 CTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
1593 CTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1652
1201 ATCCCGGCTCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
1653 ATCCCGGCTCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1712
1261 GACACACCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
1713 GACACACCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1772
1321 CTTCTGCGCGCTCTCTCGGCTACTCTATGTTTTCGTGGCGGACTACTCTTAATTTCTAGTT 1380
1773 CTTCTGCGCGCTCTCTCGGCTACTCTATGTTTTCGTGGCGGACTACTCTTAATTTCTAGTT 1832

ORIGIN		/note="Single subclone region"	
Query Match		98.5%; Score 1672.4; DB 9; Length 161835;	
Best Local Similarity		99.8%; Pred. No. 0;	
Matches 1695; Conservative		0; Mismatches 1; Indels 2; Gaps 2;	
Qy	1	AGAACCATATAGTGGTGGCCCAAGCTTATATTAATTAATTTGGCTCTCAGAAAAAGGTT	60
Db	28812	AGAACCATATAGTGGTGGCCCAAGCTTATATTAATTAATTTGGCTCTCAGAAAAAGGTT	28871
Qy	61	CCTACCCCTTGGCTCTAAGCAATTTGGTGAATGATGCGCTTACAGTAGTGACAGTAGAATGAA	120
Db	28872	CCTACCCCTTGGCTCTAAGCAATTTGGTGAATGATGCGCTTACAGTAGTGACAGTAGAATGAA	28931
Qy	121	AAGACTCTGAATTCACATTTAGAGGGTTTAGGACTGACTCAAGACGCCCTTTAGGAGGTG	180
Db	28932	AAGACTCTGAATTCACATTTAGAGGGTTTAGGACTGACTCAAGACGCCCTTTAGGAGGTG	28991
Qy	181	GAATCTACAGCCCTAGACGGCATTTGGTGGGTGAAGGAATGCTGCTTACCTCTTAATCT	240
Db	28992	GAATCTACAGCCCTAGACGGCATTTGGTGGGTGAAGGAATGCTGCTTACCTCTTAATCT	29051
Qy	241	GCTTTCTCTACTTCCATTCCTTGGCCCTCTGCAACCCATTTCTCCACTCCGACGCCATTTT	300
Db	29052	GCTTTCTCTACTTCCATTCCTTGGCCCTCTGCAACCCATTTCTCCACTCCGACGCCATTTT	29111
Qy	301	TAAAAAGATGCCCTCCCTACTTATGACTCTAAAAATGCTTCTTCTCTCTCTCTCTCTCAG	360
Db	29112	TAAAAAGATGCCCTCCCTACTTATGACTCTAAAAATGCTTCTTCTCTCTCTCTCTCTCAG	29171
Qy	361	GATATATTTCCATTAATAATACCTAGTCACTGCGCCACCTCTGCAACCCCAATGTACAT	420
Db	29172	GATATATTTCCATTAATAATACCTAGTCACTGCGCCACCTCTGCAACCCCAATGTACAT	29231
Qy	421	TCGAGTCTTACTGAACTACTTGAATTTCCCGAGATCTCACCTCTCTCGCCTGTAC	480
Db	29232	TCGAGTCTTACTGAACTACTTGAATTTCCCGAGATCTCACCTCTCTCGCCTGTAC	29291
Qy	481	CCTGTGGCGGAAAGTCAAGCCCTCCACTTCTCCCTGCTTCCACTCCCAAAATCTTCGT	540
Db	29292	CCTGTGGCGGAAAGTCAAGCCCTCCACTTCTCCCTGCTTCCACTCCCAAAATCTTCGT	29351
Qy	541	GGTTTTCAGCTCTGGAGTATTTACCGTGTGGCTGTTTAAATTTCTGCCTCCATCAGAA	600
Db	29352	GGTTTTCAGCTCTGGAGTATTTACCGTGTGGCTGTTTAAATTTCTGCCTCCATCAGAA	29411
Qy	601	GGCAGAACTGACTCGCGAACTATTCCATCCCGAGCGATAGTAGAGCTTAAAAAGAA	660
Db	29412	GGCAGAACTGACTCGCGAACTATTCCATCCCGAGCGATAGTAGAGCTTAAAAAGAA	29471
Qy	661	CGAAGAAAGTGGGTGGGAGGACTTTCAGTAACATCAGGTGGCAGCCTCAATTTTATCGTT	720
Db	29472	CGAAGAAAGTGGGTGGGAGGACTTTCAGTAACATCAGGTGGCAGCCTCAATTTTATCGTT	29530
Qy	721	TGTGAAACGTGGATAGTAATCCCTCTATCAGTGGCTGTTGCGAGGAATAAGTGAATAAA	780
Db	29531	TGTGAAACGTGGATAGTAATCCCTCTATCAGTGGCTGTTGCGAGGAATAAGTGAATAAA	29590
Qy	781	CAAAACAGGCTAGCTTGTTCATTAATGTAGTTGAATTAATCTGATTTGTTGTCAGTA	840
Db	29591	CAAAACAGGCTAGCTTGTTCATTAATGTAGTTGAATTAATCTGATTTGTTGTCAGTA	29650
Qy	841	GAAGAAGATGTGAATACCTTGAAGAAGGAGACACATTTTTTAAATATATGCTCGTAAAA	900
Db	29651	GAAGAAGATGTGAATACCTTGAAGAAGGAGACACATTTTTTAAATATATGCTCGTAAAA	29710
Qy	901	CGGATCAGAAGGCGAGGTCCCATGAGACACACCTCGCCCTTAAACATGCTCAACCCGGGC	960
Db	29711	CGGATCAGAAGGCGAGGTCCCATGAGACACACCTCGCCCTTAAACATGCTCAACCCGGGC	29770
Qy	961	TGCCATAGCTGGCGGTGCTCCCAAGGTGACTGCTCCGACAAAGGATGCTCTTCAA	1020

3398	3599	3657	3675	11517	11316
2503	2527	14158	13689	1044	1024
3341	3393	2386	2336	5448	5071
5997	5970	4939	5026	2779	2792
11401	11247	2187	2206	3013	3120
1470	1443	6457	6489	3869	3810
12564	12727	1771	1836	9708	9870
6660	6745	3788	3991	12923	12794
4555	4619	4838	4792	4563	4622
1729	1734	3160	3127	287	<800
406	<800	4088	4313	9519	9493
2213	2193	5832	6201	2229	2229
6922	6981	942	971	2610	2628
4531	4619	4341	4619	5661	5695
1381	1370	9735	9702	576	<800
1386	1370	4886	4792	820	830
1168	1168	15773	15940	3668	3654
15253	15479	398	<800	9025	8972
3613	3599	5122	5509	5092	5071
8174	8090	1336	1324	1026	1024
2074	2086	10995	10821		
714	<800	4329	4313		
6275	6239	16	<800		
3627	3599	1338	1324		
		978	971		
		2203	2206		
		2343	2336		
		489	<800		
		6300	6489		
		376	<800		
		2981	2969		
Location/Qualifiers					
1. .161835					
/organism="Homo sapiens"					
/mol_type="genomic DNA"					
/db_xref="taxon:9606"					
/chromosome="1"					
/clone="RP5-994D16"					
/clone_lib="RPC1 human PAC library 5"					
1. .4					
/note="Single subclone region"					
161830. .161835					
FEATURES					
source					
misc_feature					
misc_feature					

Db 29771 TGCCATAGCTCGTGGTCCCTCCAAAGTGACTGCTCCGACAAAAGGGTACGCTCTTCAA 29830
QY 1021 ACCCATACGTTTAAAGCAATTCAGAAAACCTCGGCTGTGCGCGACTACACGGCCATTA 1080
Db 29831 ACCCATACGTTTAAAGCAATTCAGAAAACCTCGGCTGTGCGCGACTACACGGCCATTA 29890
QY 1081 AAGAAAGACGACTCTATGCCCGCGTAATGTTCTCAGATCACAGGACGTTATTTGGAG 1140
Db 29891 AAGAAAGACGACTCTATGCCCGCGTAATGTTCTCAGATCACAGGACGTTATTTGGAG 29950
QY 1141 CTGGGAGGAGGAAGACCTTTTCTTCCGCGGGGCTAAGCGGTCTTCGAGCCCTCTTCCA 1200
Db 29951 CTGGGAGGAGGAAGACCTTTTCTTCCGCGGGGCTAAGCGGTCTTCGAGCCCTCTTCCA 30010
QY 1201 ATCCCGGCTCCGCGCGGTAACTCCCTGCCAGCGTTCGGCGGTGCTTTTTCAGCCGA 1260
Db 30011 ATCCCGGCTCCGCGCGGTAACTCCCTGCCAGCGTTCGGCGGTGCTTTTTCAGCCGA 30070
QY 1261 GACACAACTGAACTGGGGGCGCCGACCGCGGCTGCTCTGTGAAGTCAAGTTC 1320
Db 30071 GACACAACTGAACTGGGGGCGCCGACCGCGGCTGCTCTGTGAAGTCAAGTTC 30129
QY 1321 CTTCTCCCGTCTCTCGGTACTCTATGTTTTCGTGGCCGACTACTCTAATTTAGTT 1380
Db 30130 CTTCTCCCGTCTCTCGGTACTCTATGTTTTCGTGGCCGACTACTCTAATTTAGTT 30189
QY 1381 CCGGTCTCTATGCGCGCGCGGAGCAGGACGTTGTAGTGCAGTGAATAGCGGC 1440
Db 30190 CCGGTCTCTATGCGCGCGCGGAGCAGGACGTTGTAGTGCAGTGAATAGCGGC 30249
QY 1441 AAGGTCCTAATGAGATGAGGACTGATAAAATATTAGCAATAGAGCTAGGGATTGG 1500
Db 30250 AAGGTCCTAATGAGATGAGGACTGATAAAATATTAGCAATAGAGCTAGGGATTGG 30309
QY 1501 GTCAGTGGGAGATGACAGTACCACTGGCCAGTGAAACAAGCCTTAGGCGGGTCGTC 1560
Db 30310 GTCAGTGGGAGATGACAGTACCACTGGCCAGTGAAACAAGCCTTAGGCGGGTCGTC 30369
QY 1561 GTAGGCTTATCCGCTGTCCGCAATTCCTGCTAGTTTCGATCGGTAGCGGAGCGAG 1620
Db 30370 GTAGGCTTATCCGCTGTCCGCAATTCCTGCTAGTTTCGATCGGTAGCGGAGCGAG 30429
QY 1621 AGCGGACCCAGAGAGCCCTGAGCAGCCCAACCGCGCGCGGCTAGTTACCATACA 1680
Db 30430 AGCGGACCCAGAGAGCCCTGAGCAGCCCAACCGCGCGCGGCTAGTTACCATACA 30489
QY 1681 CCCCAGGAGGAGCGGAG 1698
Db 30490 CCCCAGGAGGAGCGGAG 30507

RESULT 3
HUMY292E04
LOCUS HUMY292E04 659 bp mRNA linear PRI 29-AUG-1998
DEFINITION Homo sapiens full length insert cDNA clone Y292E04.
ACCESSION AF086102
VERSION AF086102.1 GI:3483447
KEYWORDS FLI CDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 659)
AUTHORS Woesner, J., Tan, F., Marra, M., Kucaba, T., Vandell, M., Martin, J.,
Marth, G., Bowles, L., Wylie, T., Bowers, Y., Steptoe, M., Theising, B.,
Geisel, S., Allen, M., Underwood, K., Chappell, J., Person, B.,
Gibbons, M., Harvey, N., Pape, D., Chamberlain, A., Morales, R.,
Schurk, R., Ritter, E., Kohn, S., Swaller, T., Behymer, K., Hillier, L.,
Wilson, R. and Waterston, R.
Full Clone Sequencing of the Longest Available Member from Bach
Unigene Cluster
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 659)

AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (24-AUG-1998) Department of Genetics, Washington
University, 444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT SUBMITTED BY: Genome Sequencing Center
Department of Genetics
Washington University
St. Louis MO 63108, USA
http://genome.wustl.edu/gsc
mailto:est@watson.wustl.edu

NOTICE: This sequence represents the full insert of this cDNA. No
attempt has been made to verify whether this corresponds to the
full-length of the original mRNA from which it was derived. We
have tried to obtain double-stranded, or double chemistry sequence
across the entire clone, but potentially, there are areas in the
sequence where this level of coverage was not achieved.
Nevertheless, we are confident of the accuracy of this sequence as
all regions of low quality, as defined by PHRAP (P. Green, in
preparation), were visually inspected and edited accordingly. The
consensus quality values for this sequence have been submitted
separately.

The location of this clone is unknown.

FEATURES
source
1..659
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:290526"
/clone_lib="Soares_multiple_sclerosis_2NBHMS"
repeat_region
20..119
/rpt_family="L2"

ORIGIN
Query Match 26.98; Score 456.6; DB 9; Length 659;
Best Local Similarity 98.94; Pred. No. 9.2e-126;
Matches 470; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 353 CCCCTCAGGATATATTTCCCAATTAATATATACCTAAGTGAAGTCCACCTCTGCAACCCAA 412
Db 177 CACTTGAGGATATATTTCCCAATTAATATATACCTAAGTGAAGTCCACCTCTGCAACCCAA 236
QY 413 TGTCAATTCGAGTCTTATGAACTACTTGAAGTACTTCCGAGATTCCTCCTTTCTC 472
Db 237 TGTCAATTCGAGTCTTATGAACTACTTGAAGTACTTCCGAGATTCCTCCTTTCTC 296
QY 473 GCCTGTACCTGTGCGGGAAGTCAAGTCCCTTCCACCTTCTCCCTGCTTCCACTCCCAAA 532
Db 297 GCCTGTACCTGTGCGGGAAGTCAAGTCCCTTCCACCTTCTCCCTGCTTCCACTCCCAAA 356
QY 533 TACTTCGTGGTTTTCGAGCTCTGAGTATTTACCGTGTGGCTGTTTAAATTTCTGCTC 592
Db 357 TACTTCGTGGTTTTCGAGCTCTGAGTATTTACCGTGTGGCTGTTTAAATTTCTGCTC 416
QY 593 CATCAGAGGAGCAAACTGACTCGCGAACTATTCCATCCCGAGCCGATAGTAGCGCTTA 652
Db 417 CATCAGAGGAGCAAACTGACTCGCGAACTATTCCATCCCGAGCCGATAGTAGCGCTTA 476
QY 653 AAAAAGAACGGAAGAGCGGTGGGAGGACTTCAGTAAATCAGGTGGGAGCGCTCAATT 712
Db 477 AAAAAGAACGGAAGAGCGGTGGGAGGACTTCAGTAAATCAGGTGGGAGCGCTCAATT 535
QY 713 TTATCGTTTGTGAACGTTGATAGTATCCCTTATCAGCTGCTGTTGCGAGGATAAAG 772
Db 536 TTATCGTTTGTGAACGTTGATAGTATCCCTTATCAGCTGCTGTTGCGAGGATAAAG 595
QY 773 TGAATAAACAACAGAGCTAGCTTGTTCATTAATGTAGTGTGAATTAATCTGA 827
Db 596 TGAATAAACAACAGAGCTAGCTTGTTCATTAATGTAGTGTGAATTAATCTGA 650

```

RESULT 4
BX470216/c      162454 bp      DNA      linear      ROD 03-JUL-2003
LOCUS           Mouse DNA sequence from clone RP23-390M8 on chromosome 4, complete
DEFINITION
ACCESSION       BX470216
VERSION         BX470216
KEYWORDS        HTG.
SOURCE          Mus musculus (house mouse)
ORGANISM        Mus musculus
REFERENCE       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS         Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE           Pelan.S.
JOURNAL         Direct Submission
COMMENT         Submitted (03-JUL-2003) Wellcome Trust Sanger Institute, Hinxton,
                Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
                humquerry@sanger.ac.uk
                On Jul 3, 2003 this sequence version replaced gi:32439594.
                Sequence from the Mouse Genome Sequencing Consortium whole genome
                shotgun may have been used to confirm this sequence. Sequence data
                from the whole genome shotgun alone has only been used where it has
                a phred quality of at least 30.
                ----- Genome Center
                Center: Wellcome Trust Sanger Institute
                Center code: SC
                Web site: http://www.sanger.ac.uk
                Contact: humquerry@sanger.ac.uk
                -----
                During sequence assembly data is compared from overlapping clones.
                Where differences are found these are annotated as variations
                together with a note of the overlapping clone name. Note that the
                variation annotation may not be found in the sequence submission
                corresponding to the overlapping clone, as we submit sequences with
                only a small overlap as described above.
                This sequence was finished as follows unless otherwise noted: all
                regions were either double-stranded or sequenced with an alternate
                chemistry or covered by high quality data (i.e., phred quality >=
                30); an attempt was made to resolve all sequencing problems, such
                as compressions and repeats; all regions were covered by at least
                one plasmid subclone or more than one M13 subclone; and the
                assembly was confirmed by restriction digest, except on the rare
                occasion of the clone being a YAC.
                The following abbreviations are used to associate primary accession
                numbers given in the feature table with their source databases:
                Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information
                on the WORMPEP database can be found at
                http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-390M8 is
                from the RP23-23 Mouse BAC Library
                constructed by the group of Pieter de Jong.
                For further details see http://www.chori.org/bacpac/home.htm
                VECTOR: pBAC3.6.
                FEATURES
                source
                Location/Qualifiers
                1..162454
                /organism="Mus musculus"
                /mol_type="genomic DNA"
                /db_xref="taxon:10090"
                /chromosome="4"
                /clone="RP23-390M8"
                /clone_lib="RP23-23"
                ORIGIN
                Query Match      11.0%; Score 186.8; DB 10; Length 162454;
                Best Local Similarity 70.9%; Pred. No. 8.9e-44;
                Matches 346; Conservative 0; Mismatches 132; Indels 10; Gaps 7;
                Qy 1212 GCGCGGGTAATCCTCGCCAGCGTTCGGGGTGCCTTTTTCAGCGGACACACCCCT 1271
                Db 7620 GCGCGGGGTTTACCCAGCCGTGGAGAGGAGCCCTTCTAGGAGCGCGAGTAGGCGCGG 7561
                Qy 1272 GAACGTGGGGCGCCGACGCGCGCGCTGCTGGAGTACGTTCTTCTGCGCGT 1331
                Db 7560 CGATCGCGCGCCACGTGCGGACGACGCGACCCCATCTGGGAACGTTCCCTCTGCGCGT 7501

```

```

Qy 1332 CCTCTCGGCTACTCTATGTTTTCGTGGCGGACTACTCTAATTTAGTTCGGTCTCTAT 1391
Db 7500 CCTC-CGGTACTCTATGTTGCTCACGGCGA-GGCTCTCTATTTAGTTCGGTCTCTAT 7443
Qy 1392 GCGCGCGCGGAGGAGGAAACGGTTTAGTGTGACTGAATTAGCCGCCAAAGT-CCAA 1450
Db 7442 GG-----CGGCGGAGGCTGTGAGAGCTGTGGGTGCGGCGAATTAGTCTGAAGATACCA 7387
Qy 1451 TGAGNATGGAGGACTGATATAAATATTAGCAATAGAGCTAGGATGGGTTGGGTGAGTGGG 1510
Db 7386 TGAGAGCGCAGGAGCCCGCGGAGCGCTGAGCCAAATAGGCGCTAGAGA-CCCGTCAAGTGGT 7328
Qy 1511 CAGATTGACAGTACCACTGCGCCAGTGAACAACGCTAGGCGCGGTCTGTTAGGGCTTA 1570
Db 7327 GAGATGACACAGTACGCGACCAATGGGAGGAGGAGGCGGTCTATCGCGGGCTTA 7268
Qy 1571 TCCGCGCTGTCCCGCCATTCTTCGTAGTTGATCGGTAGCGGAGCGGAGCGGACCCC 1630
Db 7267 TCCGCGCGTCCCGCCATTCTTCGTAGTTGATCGGTAGCGGAGCGGAGCGGACCCC 7208
Qy 1631 AGAGAGCCCTAGCAGACGCCACCGCCCGCGCTAGTACCATCACACCCCGGAGG 1690
Db 7207 AGAGAGCCCTGAG-AGCCCCACCGCGCGCTAGTACCATCACA-CCCGGAGG 7150
Qy 1691 AGCGCGCAG 1698
Db 7149 AGCGCGCAG 7142

```

```

RESULT 5
AC101988/c      194020 bp      DNA      linear      HTG 19-MAR-2003
LOCUS           Mus musculus clone RP24-363D14, WORKING DRAFT SEQUENCE, 14
DEFINITION
ACCESSION       AC101988
VERSION         AC101988.3 GI:29124242
KEYWORDS        HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE          Mus musculus (house mouse)
ORGANISM        Mus musculus
REFERENCE       1 (bases 1 to 194020)
AUTHORS         Birren,B., Nusbaum,C. and Lander,E.
TITLE           Mus musculus, clone RP24-363D14
JOURNAL         Unpublished
REFERENCE       2 (bases 1 to 194020)
AUTHORS         Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
                Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B.,
                Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
                Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
                Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
                Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
                Ginde,S., Gird,S., Goyette,M., Graham,L., Grand-Pierre,N.,
                Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
                Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K.,
                Lamazares,R., Landers,T., Lehotzky,J., Levine,R., Liu,G.,
                MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
                McCarthy,M., McGwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
                Menes,L., Mihova,T., Menga,V., Murphy,T., Naylor,J., Nguyen,C.,
                Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
                Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
                Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
                Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
                Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
                Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
                Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
                Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
                Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
                Direct Submission
                Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
                Research, 320 Charles Street, Cambridge, MA 02141, USA
                3 (bases 1 to 194020)
                REFERENCE

```


QY 1511 CAGATTACAGTACCTGCGCAGTGAACACGCTTAGGGCGGTCTCGTAGGGCTTA 1570
 Db 120081 GAGATGACAAACGATCGGACCAATGGGGAGGACGAGGGCGGTCTATCGGGGGCTTA 120022
 QY 1571 TCCCGCTGTCGCCCGCTTCTCGTAGTTCGATCGGTAGCGGAGCGGAGCGGACGCC 1630
 Db 120021 TCCCGCGCTGCCCGCTTCTCGTAGTTCGATCGGTAGCGGAGCGGAGCGGACGCC 119862
 QY 1631 AGAGAGCCCTGAGACGCCACCGCGCGCGCGCTAGTTCACATCACACCCCGGAGG 1690
 Db 119961 AGAGAGCCCTGAG-AGCCCCACCGCGCGCGCGCTAGTTCACATCAC-CCCGGAGG 119904
 QY 1691 AGCCGCGAG 1698
 Db 119903 AGCCGCGAG 119896

RESULT 6
 AC098918/c
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-187N5, *** SEQUENCING IN PROGRESS

 AC098918 7 GI:30581384
 VERSION HTG; HTGS PHASE2; HTGS DRAFT; HTGS ENRICHED.
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 225308)

REFERENCE
 AUTHORS Muzny,D.,Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaika,D., Barber,M., Barnstead,M., Benahmed,P., Biswal,D., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Evans,K., Egan,A., Escoto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Cantar,R., Garcia,A., Garner,T., Garza,M., Gunaratne,P., Haaland,W.P., Hamil,C., Hamilton,C., Hamilton,K., Guevarra,W., Harvey,Y., Havlak,P., Haves,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hoques,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpthy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,P., London,P., Longacre,S., Lopez,J., Lorenshewa,L., Loulsged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindratne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mathew,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K., Nwaokeme,O., Okwuonu,G., Olarnpusagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C., Plopper,F., Plopper,A., Popovic,D., Primus,E., Fu,L.-L., Puzo,M., Quiroz,J., Rachin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,P., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,

Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J., Wang,S., Wang,S., Warren,J., Warren,R., Wei,X., White,P., Williams,G., Willson,R., Wlezyk,R., Woodson,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,P., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

Direct Submission
 Unpublished
 2 (bases 1 to 225308)
 Worley,K.C.
 Direct Submission
 Submitted (06-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 225308)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On May 13, 2003 this sequence version replaced gi:24942594.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GILR
 Center clone name: CH230-187N5
 ----- Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 210466 bases at least Q40
 Consensus quality: 213831 bases at least Q30
 Consensus quality: 216522 bases at least Q20
 Estimated insert size: 224754; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: this is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 225308: contig of 225308 bp in length.
 * Location/Qualifiers
 1. 225308
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-187N5"
 34231..36103
 /note="wgs_contig"

FEATURES
 source
 misc_feature
 ORIGIN

Query Match	11.0%;	Score 186.2;	DB 2;	Length 225308;
Best Local Similarity	79.7%;	Pred. No. 1.5e-43;		
Matches 307;	Conservative 0;	Mismatches 68;	Indels 10;	Gaps 7;
Qy	1315	ACGTTCCCTCTCCCGGTCTCTCGGGTACTCTATGGTTTTCGTGGCCGACTACTCTAAAT	1374	
Db	207314	ACGTTCCCTCTCGGGGTCTCTCCGGTACTCTATGGTTTTCGTGGCCCA-GGCTCTCAT	207257	
Qy	1375	CTAGTTCGGTCTCTATGCGCGCGCGGAGGCCAGGAACGGTTGTATAGTTCGACTGAATTA	1434	
Db	207256	CTAGTTCGGTCTCTATGG- - -CGCGCGAGCGCGAGAGCTGTGGTTCGGCGGAATTA	207201	
Qy	1435	GCGCCCAAGEN-CCAAATGGAATGGAGGACTGATAAAATATTAGCCATAGAGCTAGG	1493	
Db	207200	GTCCGTAAGGTACCAATAGAGAGCACAGAGCCCGCAGCGGTAGCCCAATAGTGTCTAG	207141	
Qy	1494	GATTGGGGTCAGGTGGCAGATTGACAGTACCCTGGCCAGTGAACAAAGCCCTAGGCGG	1553	
Db	207140	GA-CCCGGTCAGGTGGTGGATGGACACGATGTAGCCAAATGGAGAGGACGAGGCGG	207082	
Qy	1554	GTGCTCTGTAGGGCTTATCCCGCCTGTCTCCGCCAATTCCTCGTATGTCGATCGGTAGCGG	1613	
Db	207081	GTCTACCGGGGGCTTATCCCGCCCGTCCCGCCCAATCTCGCTAGTTTCGATCGGTAGCGG	207022	
Qy	1614	AGCGGAGAGCGGACCCGAGAGAGCCCTCAGCAGGCCCCACCGCCGCGCGCGCCCTAGTTCAC	1673	
Db	207021	AGCGGAGAGCGGACCCGAGAGAGCCCTCAG-AGCCCCACCGCCGCGCGCCCTAGTTCAC	206963	
Qy	1674	CATCACACCCCGGAGGACCGCAG	1698	
Db	206962	CATCACA-CCCGGAGGAGCGCGCAG	206939	

RESULT 7	
LOCUS	AX897942 . 178 bp DNA linear PAT 18-DEC-2003
DEFINITION	Sequence 13805 from Patent EP1033401.
ACCESSION	AX897942
VERSION	AX897942.1 GI:40052855
KEYWORDS	.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
TITLE	Expressed sequence tags and encoded human proteins
JOURNAL	Patent: EP 1033401-A 13805 06-SEP-2000; Genset (FR)

```

Contact "Location/Qualifiers
source 1..178
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 6.5%; Score 111; DB 6; Length 178;
Best Local Similarity 99.1%; Pred. No. 1.1e-21;
Matches 111; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1587 ATTCTCGCTAGTTCGATCGGTAGCGGAGCGGAGAGCCCGAGAGCCCTGAGCAG 1646
Db 1 ATTCTCGCTAGTTCGATCGGTAGCGGAGCGGAGAGCCCGAGAGCCCTGAGCAG 60

Qy 1647 CCCACCCGCCCGCGCGCTAGTTACCATCACACCCGGAGGAGCCGCAG 1698
Db 61 CCCACCCGCCCGCGCGCTAGTTACCATCACACCCGGAGGAGCCGCAG 112

```

RESULT 8			
BD033475			
LOCUS	BD033475	178 bp	DNA
DEFINITION	Sequence tag and encoded human protein.		
		linear	PAT 27-AUG-2002

BD033475
 BD033475.1 GI:22575217
 VERSION
 JP 2001269182-A/9721.
 KEYWORDS
 Homo sapiens (human)
 SOURCE
 Homo sapiens
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 178)
 REFERENCE
 Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
 AUTHORS
 Sequence tag and encoded human protein
 TITLE
 Patent: JP 2001269182-A 9721 02-OCT-2001;
 JOURNAL
 GENSET

COMMENT	GENS21
OS Homo sapiens (human)	
PN JP 2001269182-A/9721	
PD 02-OCT-2001	
PF 24-FEB-2000 JP 2000118773	
PR 26-FEB-1999 US 60/122487	
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEAN YVES JORDAN	
PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC C12N5/10,	
PC C12P21/02, C12P21/08, C12Q1/68//G06F17/30, C12N15/00, C12N5/00, PC G06F15/40	

FEATURES	source	key	Location/Qualifiers	Location/Qualifiers
			1..178	
			/organism="Homo sapiens"	
			/mol_type="genomic DNA"	
			/db_xref="taxon:9606"	

	Query Match	6.5%	Score 111;	DB 6;	Length 178;
	Best Local Similarity	99.1%;	Pred. No. 1.1e-21;		
	Matches 111;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1587	ATTCTCGTAGTTTCGATTCG	TACGGGAGCGGAGACGGAC	CCCCAGAGAGCCCTGAGCAG	1646
Db	1	ATTCTCGTAGTTTCGATTCG	TACGGGAGCGGAGACGGAC	CCCCAGAGAGCCCTGAGCAG	60
Qy	1647	CCCCAGCGCGCGCGCGCTAG	TTCACATCA	CACCCCGGAGAGCGCGAG	1698
Db	61	CCCCAGCGCGCGCGCGCTAG	TTCACATCA	CACCCCGGAGAGCGCGAG	112

RESULT 9	
AX014868	
LOCUS	1548 bp DNA
DEFINITION	Sequence 58 from Patent WO9553040.
ACCESSION	AX014868
VERSION	AX014868.1 GI:10041135
KEYWORDS	Homo sapiens (human)
SOURCE	

1
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES

Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and
Pilarczyk,C.
Human nucleic acid sequences from ovarian tumour tissue
Patent: WO 9953040-A 58 21-Oct-1999;
SCHMITT ARNO (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
(DE); PILARSKY CHRISTIAN (DE)
Location/Qualifiers
1. .1548
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Query Match 6.3%; Score 107.4; DB 6; Length 1548;
Best Local Similarity 99.1%; Pred. No. 2.4e-20;

[illegible]

Db 111282 CACCGTGGCTAGTACCATCACCCGGGAGGAGCGCAG 111240

RESULT 15
AL135841/c

LOCUS

DEFINITION Human DNA sequence from clone RP11-327L3 on chromosome 9p13.1-13.3. Contains a putative novel gene, a PGAM1 (phosphoglycerate mutase 1 (brain)) pseudogene, the gene for a novel 7 transmembrane receptor (rhodopsin family) (olfactory receptor like) protein, an NSBP1 (nuclease sensitive element binding protein 1) (YB1) pseudogene, ESTs, STSS and GSSs, complete sequence.

ACCESSION

AL135841 GI:7529194

VERSION

HTG; NSBP1; olfactory receptor; PGAM1; rhodopsin; YB1.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 99886)

Corby,N.

Direct Submission

Submitted (20-APR-2000) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

requests: clonequery@sanger.ac.uk

On Apr 9, 2000 this sequence version replaced gi:7378508.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clones, as we submit sequences with

only a small overlap as described above.

This sequence has been finished according to sequence map criteria

as follows. An attempt is made to resolve all sequencing problems,

such as compressions and repeats, but not necessarily within known

annotated human repeat sequence elements (e.g. Alu). Where the

sequence is ambiguous, there is an annotation using the 'unsure'

feature key.

The following abbreviations are used to associate primary accession

numbers given in the feature table with their source databases:

Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information

on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence

was generated from part of bacterial clone contigs of human

chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping

Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr-9

RP11-327L3 is from the library RPCI-11.2 constructed at the Roswell

Park Cancer Institute by the group of Pieter de Jong. For further

details see http://bacpac.med.buffalo.edu/

VECTOR: pBACE3.6

IMPORTANT: This sequence is not the entire insert of clone

RP11-327L3 It may be shorter because we sequence overlapping

sections only once, except for a 100 base overlap.

The true right end of clone RP11-327L3 is at 99886 in this

sequence. The true left end of clone RP11-113A10 is at 64592 in

this sequence. The true right end of clone RP11-112J3 is at 100 in

this sequence.

Location/Qualifiers

1..99886

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosomes="9"

/map="p13.1-13.3"

/clone="RP11-327L3"

/clone_lib="RPCI-11.2"

4..752

/note="LIM4c repeat: matches 1256..2036 of consensus"

repeat_region
2469..2811
/note="LIM4 repeat: matches 551..1226 of consensus"
2826..3126
/note="LIM4 repeat: matches 2333..2682 of consensus"
3139..3277
/note="AluY repeat: matches 1..301 of consensus"
3278..4323
/note="LIM4 repeat: matches 2408..2201 of consensus"
4312..4625
/note="LIM4 repeat: matches 1285..2265 of consensus"
4599..4699
/note="LIM4 repeat: matches 7467..7739 of consensus"
4730..4811
/note="LIM4 repeat: matches 1220..1315 of consensus"
4812..4921
/note="41 copies 2 mer ta 76% conserved"
4935..5210
/note="55 copies 2 mer at 71% conserved"
5250..5580
/note="AluSq repeat: matches 1..280 of consensus"
5621..5903
/note="LIM4 repeat: matches 5470..5805 of consensus"
5904..6073
/note="AluJo repeat: matches 1..282 of consensus"
6256..6828
/note="LIM4 repeat: matches 5313..5486 of consensus"
7331..7900
/note="LIM4 repeat: matches 4490..5076 of consensus"
7461..7660
/note="match: GSS: Em:AP124922"
7729..8207
/note="L1 repeat: matches 4406..4608 of consensus"
8421..8478
/note="match: GSS: Em:AQ334989"
8689..9392
/note="29 copies 2 mer ta 84% conserved"
9393..9662
/note="LIM4 repeat: matches 4081..4796 of consensus"
9663..9724
/note="AluX repeat: matches 22..236 of consensus"
9725..10042
/note="LIM4 repeat: matches 4019..4081 of consensus"
10043..10135
/note="AluYb repeat: matches 1..309 of consensus"
10303..10338
/note="LIM4 repeat: matches 3951..4019 of consensus"
10407..10730
/note="18 copies 2 mer ca 100% conserved"
10775..10919
/note="LIM4 repeat: matches 3633..3961 of consensus"
10947..10999
/note="LIM4 repeat: matches 3370..3523 of consensus"
11000..11302
/note="LIM4 repeat: matches 6121..6173 of consensus"
11303..12229
/note="AluX repeat: matches 1..311 of consensus"
12230..12522
/note="LIM4 repeat: matches 5195..6121 of consensus"
12523..12569
/note="AluSq repeat: matches 1..301 of consensus"
12570..12880
/note="LIM4 repeat: matches 5149..5195 of consensus"
12881..13168
/note="AluY repeat: matches 1..311 of consensus"
13170..13569
/note="LIM4 repeat: matches 4866..5149 of consensus"
13570..13861
/note="LIM4 repeat: matches 3212..3610 of consensus"
13862..13966
/note="AluX repeat: matches 1..310 of consensus"
14071..14391
/note="LIM4 repeat: matches 3108..3212 of consensus"
/note="AluJo repeat: matches 3..309 of consensus"

	Query Match	5.6%;	Score 95;	DB 9;	Length 99886;	
	Best Local Similarity	95.1%;	Pred. No. 3.8e-16;			
	Matches 98;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;	
Qy	1596 AGTTCGATCGGTAGCGGGACGGAGACGGACCACCCAGAGGCCCTTGACGACGCCCAACCGC	1555				
Dd	97059 AGTTCAATCGGTAGCGGGACGGAGAGCTAGCCCAGAGAGCCTTGGCACGCCCAACTC	97000				
Qy	1556 CGCGCGCGGCCTAGTTACCATCACACCCGSGAGAGCCGCAG	1698				
Dd	96999 CGCGCGCGGCCTAGTTACCATCACACCCGSGAGAGCCGCAG	96957				

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2004, 12:38:54 ; Search time 755 Seconds
(without alignments)
9554.226 Million cell updates/sec

Title: US-09-869-508-1_COPY_453_2150

Perfect score: 1698

Sequence: 1 agaacatattgtcggccaa.....caccctggaggagcgagcgcag 1698

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 29Jan04: **
1: geneseqn1980s: **
2: geneseqn1980s: **
3: geneseqn2000s: **
4: geneseqn2001as: **
5: geneseqn2001bs: **
6: geneseqn2002s: **
7: geneseqn2003as: **
8: geneseqn2003bs: **
9: geneseqn2003cs: **
10: geneseqn2004s: **

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	469.6	27.7	2102	5	AAS83003 DNA encod
2	111	6.5	178	3	AAC09730 Human sec
3	107.4	6.3	1548	2	Aaz77507 Human ova
4	95	5.6	1978	7	AAL51561 Human nuc
5	89.4	5.3	395	6	AAS61724 Lung smal
6	89.4	5.3	395	9	ADD66913 Human lun
7	89.4	5.3	395	9	ADBE8167 Human lun
8	89	5.2	1481	4	AAC81319 Human Y-b
9	89	5.2	1481	6	ABN54998 Human YB-
10	89	5.2	1481	6	ABN55602 Gene #210
11	89	5.2	1481	9	ABZ99616 Human Y-b
12	88.6	5.2	179	4	AAL02684 Human rep
13	87.4	5.1	1520	9	ADB47382 Human CDN
14	85.2	5.0	410	4	AAD23430 Human lun
15	85.2	5.0	410	9	ADD66708 Human lun
16	85.2	5.0	410	9	ADE87962 Human lun
17	83	4.9	3073	2	AAQ14635 Clone ass
18	83	4.9	3073	2	AAT34371 Plasmid p
19	83	4.9	3073	2	Aaz32246 Human gli
20	83	4.9	3073	3	AA88181 PATG29 hu
21	81.4	4.8	1474	6	ABK84103 Human cDN
22	78.8	4.6	381	3	AAC65939 Human lun
23	78.8	4.6	381	6	ABL49158 Human lun

24	78.8	4.6	381	6	ABQ92344	Abq92344 Human lun
25	78.8	4.6	381	8	ADA28305	Ada28305 Human lun
26	78.8	4.6	381	9	ADBS3510	Ades3510 Human lun
27	77.4	4.6	545	6	AAS61751	Aas61751 Lung smal
28	77.4	4.6	545	9	ADD66940	Add66940 Human lun
29	77.4	4.6	545	9	ADBE8194	Ade88194 Human lun
30	73.4	4.3	1516	7	AAD55839	Aad55839 Human nuc
31	61.2	3.6	2179	4	AAH15631	Aah15631 Human cDN
32	47.4	2.8	60	6	ABN49420	Abn49420 Human spl
33	46.4	2.7	2000	7	ADA71938	Ada71938 Rice gene
34	45.2	2.7	3900	4	AAS26731	Aas26731 Human gen
35	45.2	2.7	3900	7	ABX74080	Abx74080 Human nov
36	45.2	2.7	3901	4	AAS26729	Aas26729 Human nov
37	45.2	2.7	3901	7	ABX74078	Abx74078 Human nov
38	44.8	2.6	3152	4	AAS28912	Aas28912 Human imm
39	44.8	2.6	3152	9	ADB31753	Adb31753 Human nov
40	41.8	2.5	4140	1	AAN92735	Aan92735 Sheep PrP
41	41.6	2.4	1469	6	ABI99233	Abi99233 Mouse isc
42	41.4	2.4	10732	3	AAAI0594	Aaai0594 Gene enco
43	41.2	2.4	2452	6	AAI71445	Aai71445 TNPR/NGER
44	41	2.4	602	4	RAI37306	Rai37306 Human mus
45	41	2.4	602	7	ABX60294	Abx60294 cDNA enco

ALIGNMENTS

RESULT 1

AAS83003

ID AAS83003 standard; cDNA; 2102 BP.

AC AAS83003;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #18807.

Human; chromosome mapping; gene mapping; gene therapy; forensic;

food supplement; medical imaging; diagnostic; genetic disorder; ss.

Homo sapiens.

WO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US008631.

31-MAR-2000; 2000US-00540217.

23-AUG-2000; 2000US-00649167.

(HYSE-) HYSEQ INC.

Drmanac RT, Liu C, Tang YT;

WPI; 2001-639362/73.

P-PSDB; ABG18816.

New isolated polynucleotide and encoded polypeptides, useful in

diagnostics, forensics, gene mapping, identification of mutations

responsible for genetic disorders or other traits and to assess

biodiversity.

Claim 1; SEQ ID NO 18807; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II)

sequences. (I) is useful as hybridisation probes, polymerase chain

reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

and in recombinant production of (II). The polynucleotides are also used

in diagnostics as expressed sequence tags for identifying expressed

genes. (I) is useful in gene therapy techniques to restore normal

activity of (II) or to treat disease states involving (II). (II) is

useful for generating antibodies against it, detecting or quantitating a

PR 21-JUN-2000; 2000US-0213361P.
 PR 01-SEP-2000; 2000US-0229763P.
 PR 05-SEP-2000; 2000US-0230629P.
 PR 14-SEP-2000; 2000US-0232565P.
 PR 19-DEC-2000; 2000US-0257037P.
 PR 08-JAN-2001; 2001US-0260796P.
 XX (CORI-) CORIXA CORP.
 PA
 XX
 PI Lodes MJ, Wang T, Mohamath R, Indirias CY;
 XX
 DR WPI; 2002-010896/01.
 XX
 XX Lung tumor polynucleotide and polypeptides useful in therapy and
 PT diagnosis of cancer especially lung cancer.
 XX
 XX Claim 1; Page 226; 295pp; English.
 XX
 CC The invention relates to novel isolated lung small cell cancer antigen
 CC polynucleotides (I) and polypeptides (II) used in a method of detecting
 CC cancer in a patient. The method is optionally performed by utilising
 CC oligonucleotides (III), where the biological sample from the patient is
 CC contacted with (III), detecting the amount of polynucleotide hybridised
 CC to (III) in the sample and comparing the amount of polynucleotide to a
 CC predetermined cut-off value and thereby determining cancer in a patient.
 CC (I), (II) or antigen-presenting cells expressing (II) is useful for
 CC stimulating and/or expanding T cells specific for a tumour protein. The
 CC method comprises contacting T cells with one of the components under
 CC conditions to permit the stimulation and/or expansion of the cells. A
 CC composition comprising (I) is useful for stimulating an immune response
 CC in a patient and for inhibiting the development of a cancer especially
 CC lung cancer in a patient. An isolated T cell population is useful for
 CC removing tumour cells from the biological sample and for inhibiting the
 CC development of cancer in a patient. AAS61460-AAS61874 represent novel
 CC human lung small cell cancer antigen coding sequences of the invention
 XX
 XX Sequence 395 BP; 86 A; 138 C; 126 G; 44 T; 0 U; 1 Other;
 SQ
 Query Match 5.3%; Score 89.4; DB 6; Length 395;
 Best Local Similarity 98.9%; Pred. No. 2.2e-17;
 Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1608 AGCGGAGCGGAGAGCGGACCCAGAGAGCCCTGAGAGAGCCCGCCGCGCGCT 1667
 Db 13 AGGGGAGCGGAGAGCGGACCCAGAGAGCCCTGAGAGAGCCCGCCGCGCGCT 72
 QY 1668 AGTTACCATCACACCCCGGAGAGCGCGCAG 1698
 Db 73 AGTTACCATCACACCCCGGAGAGCGCGCAG 103
 RESULT 6
 ADD66913
 ID ADD66913 standard; cDNA; 395 BP.
 XX
 XX ADD66913;
 XX
 XX 15-JAN-2004 (first entry)
 DT
 XX Human lung tumour-specific related cDNA, SEQ ID No 605.
 DE
 XX expression control; cancer; T cell; tumour; immune; cytostatic; vaccine;
 KW human; lung tumour-specific; ss.
 XX
 XX Homo sapiens.
 OS
 XX WO200292001-A2.
 PN
 XX 21-NOV-2002.
 FD
 XX 10-MAY-2002; 2002WO-US014975.
 XX
 XX 11-MAY-2001; 2001US-00854133.
 PR

XX (CORI-) CORIXA CORP.
 PA
 XX Lodes MJ, Wang T, Fan L, Algate PA, Mcneill PD;
 XX
 XX WPI; 2003-120592/11.
 DR
 XX
 XX New polynucleotide and polypeptide, useful for preparing a composition
 PT for diagnosing, treating or preventing cancer.
 XX
 XX Claim 11; SEQ ID NO 605; 494pp; English.
 PS
 XX
 CC The invention relates to a novel isolated polynucleotide comprising one
 CC of 32 47-6080 base pair sequences, given in the specification, or their
 CC complements or degenerate variants, at least 20 contiguous residues of a
 CC sequence in, or having at least 75 or 90 % identity with the isolated
 CC polynucleotide, or that hybridise with the polynucleotide. The invention
 CC further comprises: an isolated polypeptide; an expression vector
 CC comprising the polynucleotide operably linked to an expression control
 CC sequence; a host cell transformed or transfected with the expression
 CC vector; an isolated antibody or its antigen-binding fragment that
 CC specifically binds to the polypeptide; a method for detecting the
 CC presence of a cancer in a patient; a fusion protein comprising the
 CC polypeptide; an oligonucleotide that hybridises to the isolated
 CC polynucleotide under moderately stringent conditions; a method for
 CC stimulating and/or expanding T cells specific for a tumour protein; an
 CC isolated T cell population; a composition comprising a first component
 CC consisting of carriers and immunostimulants and a second component; a
 CC method for stimulating an immune response in a patient; a method for
 CC treating cancer in a patient; a method for determining cancer in a
 CC patient; a diagnostic kit comprising at least one oligonucleotide or
 CC antibody and a detection reagent comprising a reporter group; and a
 CC method for inhibiting the development of cancer in a patient. The
 CC compositions of the invention have cytostatic activity and can be used to
 CC create a vaccine. The isolated polynucleotide is useful for preparing a
 CC composition for diagnosing, treating or preventing cancer. This
 CC polynucleotide sequence represents a human lung tumour-specific cDNA
 CC sequence relating to the invention.
 XX
 XX Sequence 395 BP; 86 A; 138 C; 126 G; 44 T; 0 U; 1 Other;
 SQ
 Query Match 5.3%; Score 89.4; DB 9; Length 395;
 Best Local Similarity 98.9%; Pred. No. 2.2e-17;
 Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1608 AGCGGAGCGGAGAGCGGACCCAGAGAGCCCTGAGAGAGCCCGCCGCGCGCT 1667
 Db 13 AGGGGAGCGGAGAGCGGACCCAGAGAGCCCTGAGAGAGCCCGCCGCGCGCT 72
 QY 1668 AGTTACCATCACACCCCGGAGAGCGCGCAG 1698
 Db 73 AGTTACCATCACACCCCGGAGAGCGCGCAG 103
 RESULT 7
 ADE88167
 ID ADE88167 standard; cDNA; 395 BP.
 XX
 XX ADE88167;
 XX
 XX 29-JAN-2004 (first entry)
 DT
 XX Human lung tumour antigen cDNA #500.
 DE
 XX Human; lung tumour antigen; gene; ss; cancer; lung cancer; CD4+; CD8+;
 KW T cell; immune response; immunostimulant; cytostatic.
 XX
 XX Homo sapiens.
 OS
 XX US2003118599-A1.
 PN
 XX 26-JUN-2003.
 XX

QY	1610	CGGAGCGGAGAGCGGAGACCCGAGAGAGCCCTGAGACAGCCCAACGCGCGCGCGCGGCGCTAG	1669
Db	2	CGGAGCGGAGAGCGGAGACCCGAGAGAGCCCTGAGACAGCCCAACGCGCGCGCGCGGCGCTAG	61
QY	1670	TTACCATCACACCCCGGAGAGCGGAGCGGAG	1698
Db	62	TTACCATCACACCCCGGAGAGCGGAGCGGAG	90
RESULT 9			
ID	ABS54998 standard; DNA; 1481 BP.		
XX	ABS54998;		
XX	10-DEC-2002 (first entry)		
XX	Human YB-1 Y Box Binding Protein gene.		
DE	Telomerase reverse transcriptase; TERT; replication-conditional virus;		
XX	adenovirus replication gene; cancer cell; lung; pancreatic cancer;		
KW	medulloblastoma; cervical carcinoma; fibrosarcoma; osteosarcoma;		
KW	cytolysis; replication defective adenovirus vector; congenital defect;		
KW	proinflammatory; antiinflammatory; heterologous effector gene;		
KW	cancer therapy; cytostatic; gene therapy; human; YB-1;		
XX	Y Box Binding Protein; gene; ds.		
XX	Homo sapiens.		
XX	OS		
PH	Key Location/Qualifiers		
FT	127..1080		
FT	/*tag= a		
FT	/product= "Human YB-1 Y Box Binding Protein"		
XX	WO200253760-A2.		
PN	XX		
XX	11-JUL-2002.		
XX	XX		
XX	17-DEC-2001; 2001WO-US048785.		
XX	XX		
PR	18-DEC-2000; 2000US-0256418P.		
XX	XX		
XX	(GERO-) GERON CORP.		
PA	XX		
PI	Irving JM, Lebkowski JS;		
XX	XX		
XX	WPI; 2002-732123/78.		
DR	P-FSDS; ABG70989.		
XX	XX		
PT	Novel replication-conditional virus useful for cytolysis of target cells		
PT	e.g. cancer cells and preparing a medicament for treating cancer,		
PT	comprises heterologous replication element in an adenovirus-based		
XX	construct.		
XX	Disclosure; Page 30; 32pp; English.		
PS	XX		
XX	The present invention relates to a new replication-conditional virus with		
CC	a genome inventing adenovirus replication genes and one or more		
CC	heterologous gene(s) that functionally replaces one or more adenovirus		
CC	gene(s) required for replication or assembly of the virus. The invention		
CC	is useful for killing a cancer cell (such as lung, pancreatic cancer,		
CC	medulloblastoma, cervical carcinoma, fibrosarcoma or osteosarcoma),		
CC	killing a cell expressing TERT (telomerase reverse transcriptase), and in		
CC	preparing a medicament for treating cancer and a condition associated		
CC	with increased expression of TERT in affected cells, in a subject. The		
CC	invention is also useful for cytolysis of specific target cells. The		
CC	invention is further useful for producing replication defective		
CC	adenovirus vector which is useful for transient expression of a		
CC	heterologous therapeutic gene to correct a congenital defect, introducing		
CC	proinflammatory or antiinflammatory activity, enhancing telomerase		
CC	function, and delivering heterologous effector genes that induce killing		
CC	of the transduced cells. The invention is more safe for use in cancer		
CC	therapy. The present nucleic acid sequence encodes the human YB-1 Y Box		

CC Binding Protein as described in the invention

XX
SQ Sequence 1481 BP; 437 A; 381 C; 386 G; 277 T; 0 U; 0 Other;

Query Match 5.2%; Score 89; DB 6; Length 1481;
Best Local Similarity 100.0%; Pred. No. 7.1e-17;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1610 CGGAGCGGAGAGCGGACCCCGAGAGAGCCCTGAGAGCCCGACCGCCCGCGCGGCTAG 1669
DB 2 CGGAGCGGAGAGCGGACCCCGAGAGAGCCCTGAGAGCCCGACCGCCCGCGCGGCTAG 61

QY 1670 TTACCATCACACCCCGGAGAGCCCGAG 1698
DB 62 TTACCATCACACCCCGGAGAGCCCGAG 90

RESULT 10

ABN95602

ID ABN95602 standard; DNA; 1481 BP.

XX
AC ABN95602;

DT 13-AUG-2002 (first entry)

XX
DE Gene #2100 used to diagnose liver cancer.

XX
KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.

XX
OS Homo sapiens.

XX
PN WO200229103-A2.

XX
PD 11-APR-2002.

XX
PF 02-OCT-2001; 2001WO-US030589.

XX
PR 02-OCT-2000; 2000US-0237054P.

XX
PA (GENE-) GENE LOGIC INC.

XX
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;

XX
DR WPI; 2002-426119/45.

XX
PT Diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma or metastatic liver tumor in a patient, involves detecting the level of expression of two or more genes in a liver tissue sample.

XX
PS Claim 1; SEQ ID NO 2100; 298pp; English.

XX
CC The invention relates to a novel method for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma or metastatic liver tumor in a patient, and differentiating metastatic liver cancer from hepatocellular carcinoma in a patient, involving detecting the level of expression of two or more genes represented in ABN9503-ABN97455 in a tissue sample. The method of the invention has hepatotropic, and cytosstatic activity. The method is useful for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma and metastatic liver carcinoma in a patient. The method is useful for identifying expression profiles which serve as useful diagnostic markers as well as markers that can be used to monitor disease states, disease progression, drug toxicity, drug efficacy and drug metabolism. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

XX
SQ Sequence 1481 BP; 437 A; 381 C; 386 G; 277 T; 0 U; 0 Other;

Query Match 5.2%; Score 89; DB 6; Length 1481;
Best Local Similarity 100.0%; Pred. No. 7.1e-17;

PR	06-SEP-2000;	2000US-0230438P.
PR	07-SEP-2000;	2000US-02311242P.
PR	08-SEP-2000;	2000US-0231243P.
PR	09-SEP-2000;	2000US-0231244P.
PR	08-SEP-2000;	2000US-0231413P.
PR	08-SEP-2000;	2000US-0231414P.
PR	08-SEP-2000;	2000US-0232080P.
PR	08-SEP-2000;	2000US-0232081P.
PR	12-SEP-2000;	2000US-0231968P.
PR	14-SEP-2000;	2000US-0232357P.
PR	14-SEP-2000;	2000US-0232358P.
PR	14-SEP-2000;	2000US-0232359P.
PR	14-SEP-2000;	2000US-0232401P.
PR	14-SEP-2000;	2000US-0233063P.
PR	14-SEP-2000;	2000US-0233064P.
PR	14-SEP-2000;	2000US-0233065P.
PR	21-SEP-2000;	2000US-0234223P.
PR	21-SEP-2000;	2000US-0234274P.
PR	25-SEP-2000;	2000US-0234997P.
PR	25-SEP-2000;	2000US-0235498P.
PR	26-SEP-2000;	2000US-0235484P.
PR	27-SEP-2000;	2000US-0235834P.
PR	27-SEP-2000;	2000US-0235836P.
PR	29-SEP-2000;	2000US-0236327P.
PR	29-SEP-2000;	2000US-0236367P.
PR	29-SEP-2000;	2000US-0236368P.
PR	29-SEP-2000;	2000US-0236369P.
PR	29-SEP-2000;	2000US-0236370P.
PR	02-OCT-2000;	2000US-0236802P.
PR	02-OCT-2000;	2000US-0237037P.
PR	02-OCT-2000;	2000US-0237038P.
PR	02-OCT-2000;	2000US-0237039P.
PR	02-OCT-2000;	2000US-0237040P.
PR	13-OCT-2000;	2000US-0239935P.
PR	13-OCT-2000;	2000US-0239937P.
PR	20-OCT-2000;	2000US-0240960P.
PR	20-OCT-2000;	2000US-0241212P.
PR	20-OCT-2000;	2000US-0241785P.
PR	20-OCT-2000;	2000US-0241786P.
PR	20-OCT-2000;	2000US-0241808P.
PR	20-OCT-2000;	2000US-0241809P.
PR	20-OCT-2000;	2000US-0241826P.
PR	01-NOV-2000;	2000US-0246173P.
PR	08-NOV-2000;	2000US-0246474P.
PR	08-NOV-2000;	2000US-0246475P.
PR	08-NOV-2000;	2000US-0246476P.
PR	08-NOV-2000;	2000US-0246477P.
PR	08-NOV-2000;	2000US-0246478P.
PR	08-NOV-2000;	2000US-0246523P.
PR	08-NOV-2000;	2000US-0246524P.
PR	08-NOV-2000;	2000US-0246525P.
PR	08-NOV-2000;	2000US-0246526P.
PR	08-NOV-2000;	2000US-0246527P.
PR	08-NOV-2000;	2000US-0246528P.
PR	08-NOV-2000;	2000US-0246529P.
PR	17-NOV-2000;	2000US-0249208P.
PR	17-NOV-2000;	2000US-0249209P.
PR	17-NOV-2000;	2000US-0249210P.
PR	17-NOV-2000;	2000US-0249211P.
PR	17-NOV-2000;	2000US-0249212P.
PR	17-NOV-2000;	2000US-0249213P.
PR	17-NOV-2000;	2000US-0249214P.
PR	17-NOV-2000;	2000US-0249215P.
PR	17-NOV-2000;	2000US-0249216P.
PR	17-NOV-2000;	2000US-0249217P.
PR	17-NOV-2000;	2000US-0249218P.

PF 03-OCT-2001; 2001US-00971392.
XX
XX
PR 03-OCT-2000; 2000US-0237652P.
XX
XX (PETE/) PETERSON D P.
PA (PEAR/) PEARSON C I.
PA (COCK/) COCKS B G.
XX
XX Peterson DP, Pearson CI, Cocks BG;
XX WPI; 2003-662509/62.
XX
XX New combination comprises cDNAs that are differentially expressed in
PT dendritic cells useful for preparing a composition for diagnosing or
PT treating cancer, infectious disease, autoimmunity, allergy or graft
PT versus host disease.
XX
XX Claim 1; SEQ ID NO 82; 28pp; English.
XX
XX The invention relates to a combination comprising cDNAs that are
CC differentially expressed in dendritic cells (DC). Also included is a high
CC throughput method for detecting differential expression of one or more
CC cDNAs in a sample containing nucleic acids. The combination is useful for
CC preparing a composition for diagnosing, treating and monitoring the
CC treatment of cancer, infectious disease, autoimmunity, allergy or graft
CC versus host disease, or for enhancing a vaccine. The present sequence
CC represents a human cDNA upregulated in dendritic cells. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=20030134283.
XX
XX Sequence 1520 BP; 445 A; 385 C; 393 G; 297 T; 0 U; 0 Other;
SQ
Query Match 5.1%; Score 87.4; DB 9; Length 1520;
Best Local Similarity 98.0%; Pred. No. 2.3e-16;
Matches 99; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1599 TCGATCGGTAGCGGAGCGGAGCGGACCCAGAGAGCCCTGAGCAGCCACCGCGC 1658
DB 5 TAGATCGGTAGCGGAGCGGAGCGGACCCAGAGAGCCCTGAGCAGCCACCGCGC 64
QY 1659 CGCCGGCCCTAGTTACCATCACACCCCGG-GAGGAGCGCGAG 1698
DB 65 CGCCGGCCCTAGTTACCATCACACCCCGGTTGAGGAGCGCGAG 105
RESULT 14
AAD23430
ID AAD23430 standard; cDNA; 410 BP.
AC AAD23430;
XX
XX 26-FEB-2002 (first entry)
XX
XX Human lung tumour-specific 54243.1 cDNA.
XX
XX Human; lung tumour protein; immunostimulant; cytostatic; gene therapy;
KW antisense-therapy; vaccine; immune response; lung cancer; 54243.1; ss.
XX
XX Homo sapiens.
XX
XX WO200172295-A2.
XX
XX 04-OCT-2001.
XX
XX 28-MAR-2001; 2001WO-US009991.
XX
XX 29-MAR-2000; 2000US-00538037.
PR 05-JUN-2000; 2000US-00589937.
PR 18-AUG-2000; 2000US-00640878.
PR 22-SEP-2000; 2000US-02345172.
PR 01-NOV-2000; 2000US-00704512.
PR 14-DEC-2000; 2000US-00738973.
XX

XX (CORI-) CORIXA CORP.
XX
XX Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indrias CY;
PI Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos WD;
XX
XX WPI; 2001-639201/73.
DR
XX
XX New human lung-specific polynucleotides and polypeptides for the
PT diagnosis and treatment of disease e.g. lung cancer.
XX
XX Claim 1; Page 299; 378pp; English.
XX
XX The invention relates to isolated lung tumour-specific proteins and their
CC corresponding cDNA molecules. Lung tumour-specific proteins and their
CC antigen-presenting cells are useful for stimulating and/or expanding T
CC cells specific for a tumour protein, and for inhibiting the development
CC of cancer. The invention also relates to a composition useful for
CC stimulating an immune response, and for treating cancer. The lung tumour
CC specific oligonucleotide is useful in gene therapy and for diagnosis,
CC detection and treatment of lung cancer. The present sequence is a cDNA
CC encoding human lung tumour-specific protein
XX
XX Sequence 410 BP; 97 A; 145 C; 125 G; 43 T; 0 U; 0 Other;
SQ
Query Match 5.0%; Score 85.2; DB 4; Length 410;
Best Local Similarity 96.7%; Pred. No. 5e-16;
Matches 87; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1609 GCGGAGCGGAGCGGAGCGGACCCAGAGAGCCCTGAGCAGCCACCGCGCGCCTA 1668
DB 2 GCACGAGGCGGAGCGGAGCGGACCCAGAGAGCCCTGAGCAGCCACCGCGCGCCTA 61
QY 1669 GTTACATCATCACCCCGGAGGAGCGCGAG 1698
DB 62 GTTACATCATCACCCCGGAGGAGCGCGAG 91
RESULT 15
ADD66708
ID ADD66708 standard; cDNA; 410 BP.
XX
XX ADD66708;
XX
XX 15-JAN-2004 (first entry)
XX
XX Human lung tumour-specific related cDNA, SEQ ID No 400.
DE
XX
XX expression control; cancer; T cell; tumour; immune; cytostatic; vaccine;
KW human; lung tumour-specific; ss.
XX
XX Homo sapiens.
XX
XX WO200292001-A2.
XX
XX 21-NOV-2002.
PD
XX
XX 10-MAY-2002; 2002WO-US014975.
XX
XX 11-MAY-2001; 2001US-00854133.
XX
XX (CORI-) CORIXA CORP.
XX
XX Lodes MJ, Wang T, Fan L, Algate PA, McNeill PD;
XX
XX WPI; 2003-120592/11.
XX
XX New polynucleotide and polypeptide, useful for preparing a composition
PT for diagnosing, treating or preventing cancer.
XX
XX Disclosure; SEQ ID NO 400; 494pp; English.
PS
XX
XX The invention relates to a novel isolated polynucleotide comprising one
CC

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2004, 13:37:45 ; Search time 155 Seconds
(without alignments)
6079.398 Million call updates/sec

Title: US-09-869-508-1_COPY_453_2150

Perfect score: 1698

Sequence: 1 agaacatattgtcgcccaaa.....caccocggaggagccgcag 1698

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
- 2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
- 3: /cgn2_6/prodata/2/ina/6A_COMB.seq:*
- 4: /cgn2_6/prodata/2/ina/6B_COMB.seq:*
- 5: /cgn2_6/prodata/2/ina/PTUS_COMB.seq:*
- 6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	89	5.2	1481	3	US-09-429-323-3
2	83	4.9	3073	1	US-07-688-352C-31
3	83	4.9	3073	2	US-08-474-379C-31
4	83	4.9	3073	3	US-09-146-249A-31
5	83	4.9	3073	3	US-08-206-188B-31
6	83	4.9	3073	5	PCT-US91-02714-30
7	78.8	4.6	381	4	US-09-643-597-215
8	78.8	4.6	381	4	US-09-480-884A-215
9	78.8	4.6	381	4	US-09-542-615A-215
10	78.8	4.6	381	4	US-09-606-421B-215
11	76.6	4.5	7218	1	US-08-232-463-14
12	44.2	2.6	7218	1	US-08-232-463-14
13	43.4	2.6	289	3	US-09-007-005-17
14	43.4	2.6	289	3	US-09-244-796-17
15	40.8	2.4	13104	3	US-08-256-799-4
16	40.8	2.4	13104	3	US-08-462-437-4
17	37.8	2.2	832	4	US-09-621-976-2813
18	37	2.2	505	4	US-09-621-976-2813
19	36	2.1	229	4	US-09-702-705-246
20	36	2.1	229	4	US-09-736-457-246
21	36	2.1	229	4	US-09-614-124B-246
22	36	2.1	229	4	US-09-671-325-246
23	36	2.1	229	4	US-09-589-184-246
24	36	2.1	230	4	US-09-702-705-261
25	36	2.1	230	4	US-09-736-457-261
26	36	2.1	230	4	US-09-614-124B-261
27	36	2.1	230	4	US-09-671-325-261

```

28 36 2.1 230 4 US-09-589-184-261 Sequence 261, Appl
c 29 36 2.1 472 4 US-09-976-594-87 Sequence 87, Appl
c 30 36 2.1 19513 4 US-10-204-708-39 Sequence 39, Appl
c 31 35.2 2.1 690 4 US-09-107-532A-2410 Sequence 2410, Ap
c 32 35.2 2.1 939 4 US-09-671-317-155 Sequence 155, App
c 33 35.2 2.1 1000 4 US-09-671-317-156 Sequence 156, App
c 34 35.2 2.1 1173 4 US-09-252-991A-14724 Sequence 14724, A
c 35 35.2 2.1 1686 4 US-09-252-991A-14463 Sequence 14463, A
c 36 35.2 2.1 99916 4 US-09-816-095-3 Sequence 3, Appli
c 37 35.2 2.1 392000 4 US-10-027-983-11 Sequence 11, Appli
c 38 34.8 2.0 72928 3 US-09-009-913-1 Sequence 1, Appli
c 39 34.8 2.0 640681 4 US-09-790-988-1 Sequence 1, Appli
c 40 34.6 2.0 19250 4 US-08-961-527-35 Sequence 35, Appl
c 41 34.2 2.0 98844 4 US-09-791-211-10 Sequence 10, Appl
c 42 34 2.0 498 4 US-09-621-976-18590 Sequence 18590, A
c 43 34 2.0 6305 3 US-08-687-580B-6 Sequence 6, Appli
c 44 33.8 2.0 520 4 US-09-621-976-17316 Sequence 17316, A
c 45 33.8 2.0 587 3 US-08-577-081A-9 Sequence 9, Appli

```

ALIGNMENTS

RESULT 1

```

US-09-429-323-3
; Sequence 3, Application US/09429323A
; Patent No. 6140126
; Patent No. 6140126 6140123
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Lex M. Cowsett
; TITLE OF INVENTION: ANTISENSE MODULATION OF Y-BOX BINDING PROTEIN 1 EXPRESSION
; FILE REFERENCE: RYS-0092
; CURRENT APPLICATION NUMBER: US/09/429,323A
; CURRENT FILING DATE: 1999-10-26
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 3
; LENGTH: 1481
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (127)..(1080)
US-09-429-323-3

```

```

Query Match 5.2%; Score 89; DB 3; Length 1481;
Best Local Similarity 100.0%; Pred. No. 2.7e-18;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1610 CGGAGCGGAGAGCGGACCCAGAGAGCCCTGTAGAGCCACCGCCGCGCGGCTAG 1669
Db 2 CGGAGCGGAGAGCGGACCCAGAGAGCCCTGTAGAGCCACCGCCGCGCGGCTAG 61

```

```

Qy 1670 TTACCATCACACCCCGGAGAGCCGCGAG 1698
Db 62 TTACCATCACACCCCGGAGAGCCGCGAG 90

```

RESULT 2

```

US-07-688-352C-31
; Sequence 31, Application US/07688352C
; Patent No. 5527896
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: Cloning by Complementation and Related
; TITLE OF INVENTION: Processes
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street

```

;; CITY: Chicago
;; STATE: Illinois
;; COUNTRY: USA
;; ZIP: 60603
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/07/688,352C
;; FILING DATE: 19910419
;; CLASSIFICATION: 435
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/511,715
;; FILING DATE: 20-APR-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Borun, Michael P.
;; REGISTRATION NUMBER: 25447
;; REFERENCE/DOCKET NUMBER: 27805/30197
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (312) 346-5750
;; TELEFAX: (312) 984-9740
;; TELEX: 25-3856
;; INFORMATION FOR SEQ ID NO: 31:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 3073 base pairs
;; TYPE: NUCLEIC ACID
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 3..1111
;;
US-07-688-352C-31

Query Match 4.9%; Score 83; DB 1; Length 3073;
Best Local Similarity 100.0%; Pred. No. 4.1e-16;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1616 CGGAGAGCGGACCCCGAGGAGCCCTGAGCAGCCCGCCGCGCGCGCTAGTTACCA 1675
DB 19 CGGAGAGCGGACCCCGAGGAGCCCTGAGCAGCCCGCCGCGCGCGCTAGTTACCA 78

QY 1676 TCACACCCCGGAGGAGCCCGAG 1698
DB 79 TCACACCCCGGAGGAGCCCGAG 101

RESULT 3
US-08-474-379C-31
; Sequence 31, Application US/08474379C
; Patent No. 5977305
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: CLONING BY COMPLEMENTATION AND RELATED
; TITLE OF INVENTION: PROCESSES
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,379C

;; CITY: Chicago
;; STATE: Illinois
;; COUNTRY: USA
;; ZIP: 60603
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/07/688,352C
;; FILING DATE: 19910419
;; CLASSIFICATION: 435
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/511,715
;; FILING DATE: 20-APR-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Borun, Michael P.
;; REGISTRATION NUMBER: 25447
;; REFERENCE/DOCKET NUMBER: 27805/30197
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (312) 346-5750
;; TELEFAX: (312) 984-9740
;; TELEX: 25-3856
;; INFORMATION FOR SEQ ID NO: 31:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 3073 base pairs
;; TYPE: NUCLEIC ACID
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 3..1111
;;
US-07-688-352C-31

Query Match 4.9%; Score 83; DB 1; Length 3073;
Best Local Similarity 100.0%; Pred. No. 4.1e-16;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1616 CGGAGAGCGGACCCCGAGGAGCCCTGAGCAGCCCGCCGCGCGCGCTAGTTACCA 1675
DB 19 CGGAGAGCGGACCCCGAGGAGCCCTGAGCAGCCCGCCGCGCGCGCTAGTTACCA 78

QY 1676 TCACACCCCGGAGGAGCCCGAG 1698
DB 79 TCACACCCCGGAGGAGCCCGAG 101

RESULT 4
US-09-146-249A-31
; Sequence 31, Application US/09146249A
; Patent No. 6069240
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: Cloning by Complementation and Related
; TITLE OF INVENTION: Processes
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,249A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/511,715
; FILING DATE: 20-APR-1990
; ATTORNEY/AGENT INFORMATION:

```

; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3073 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..1111
US-09-146-249A-31

Query Match
Best Local Similarity 100.0%; Pred. No. 4.1e-16; Length 3073;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..1111
US-08-206-188B-31

Query Match
Best Local Similarity 100.0%; Pred. No. 4.1e-16; Length 3073;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1616 CGGAGAGCGGACCCCGAGAGCGCTTGAGCAGCCCGCCGCGCGCTAGTTACCA 1675
Db 19 CGGAGAGCGGACCCCGAGAGCGCTTGAGCAGCCCGCCGCGCGCTAGTTACCA 78

QY 1676 TCACACCCCGGAGAGCGCGAG 1698
Db 79 TCACACCCCGGAGAGCGCGAG 101

RESULT 5
PCT-US91-02714-30
; Sequence 31, Application US/08206188B
; Patent No. 6100025
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: Cloning by Complementation and Related
; NUMBER OF SEQUENCES: 84
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/206,188B
; FILING DATE: 01-MAR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/511,715
; FILING DATE: 20-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25447
; REFERENCE/DOCKET NUMBER: 27805/30197
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3073 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..1111
PCT-US91-02714-30

Query Match
Best Local Similarity 100.0%; Pred. No. 4.1e-16; Length 3073;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1616 CGGAGAGCGGACCCCGAGAGCGCTTGAGCAGCCCGCCGCGCGCTAGTTACCA 1675
Db 19 CGGAGAGCGGACCCCGAGAGCGCTTGAGCAGCCCGCCGCGCGCTAGTTACCA 78

QY 1676 TCACACCCCGGAGAGCGCGAG 1698
Db 79 TCACACCCCGGAGAGCGCGAG 101

RESULT 6
PCT-US91-02714-30
; Sequence 30, Application PC/TUS9102714
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: Cloning by Complementation and Related
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/02714
; FILING DATE: 19910419
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/511,715
; FILING DATE: 20-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25447
; REFERENCE/DOCKET NUMBER: 27805/30197
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3073 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..1111
PCT-US91-02714-30

Query Match
Best Local Similarity 100.0%; Pred. No. 4.1e-16; Length 3073;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1616 CGGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCGCCAGCCCGCGGCTAGTTACCA 1675
Db 19 CGGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCGCCAGCCCGCGGCTAGTTACCA 78
QY 1676 TCACACCCCGGAGGAGCGCCGAG 1698
Db 79 TCACACCCCGGAGGAGCGCCGAG 101

RESULT 7
US-09-643-597-215
; Sequence 215, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy R.
; APPLICANT: Fang, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 215
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (17)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (20)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (60)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (61)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (365)
; OTHER INFORMATION: n=A,T,C or G
US-09-643-597-215

Query Match 4.6%; Score 78.8; DB 4; Length 381;
Best Local Similarity 93.0%; Pred. No. 2.3e-15;
Matches 80; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1613 GAGCGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCGCCAGCCCGCGGCTAGTTA 1672
Db 1 GAGCGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCGCCAGCCCGCGGCTAGTTA 60
QY 1673 CCATCACACCCCGGAGGAGCCGAG 1698
Db 61 NCATCACACCCCGGAGGAGCCGAG 86

RESULT 8
US-09-480-884A-215
; Sequence 215, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
```

```
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fang, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 215
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(381)
; OTHER INFORMATION: n=A,T,C or G
US-09-480-884A-215

Query Match 4.6%; Score 78.8; DB 4; Length 381;
Best Local Similarity 93.0%; Pred. No. 2.3e-15;
Matches 80; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1613 GAGCGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCGCCAGCCCGCGGCTAGTTA 1672
Db 1 GAGCGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCGCCAGCCCGCGGCTAGTTA 60
QY 1673 CCATCACACCCCGGAGGAGCCGAG 1698
Db 61 NCATCACACCCCGGAGGAGCCGAG 86

RESULT 9
US-09-542-615A-215
; Sequence 215, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fang, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 215
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (17)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (20)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (60)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (61)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (365)
; OTHER INFORMATION: n=A,T,C or G
US-09-542-615A-215
```


Db73AGTTACCATCACACCCCGGAGGCGCGCAG103

RESULT 2

US-09-854-133-605

Sequence 605, Application US/09854133

Publication No. US20020183499A1

GENERAL INFORMATION:

APPLICANT: Lodes, Michael J.

APPLICANT: Mchamath, Raodoh

APPLICANT: Henderson, Robert A.

APPLICANT: Benson, Darin R.

APPLICANT: Secrist, Heather

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR

FILE REFERENCE: 210121.475C10

CURRENT APPLICATION NUMBER: US/09/854,133

CURRENT FILING DATE: 2001-05-11

NUMBER OF SEQ ID NOS: 735

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 605

LENGTH: 395

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

LOCATION: 375

OTHER INFORMATION: n = A,T,C or G

US-09-854-133-605

Query Match5.3%; Score 89.4; DB 9; Length 395;

Best Local Similarity98.9%; Pred. No. 8.7e-18;

Matches90; Conservative0; Mismatches1; Indels0; Gaps0;

QY1608AGCGGAGCGGAGCGGACCCCGAGAGCCCTTGAGAGCCCTTGAGAGCCCGCGCGGCT1667

Db13AGCGGAGCGGAGCGGACCCCGAGAGCCCTTGAGAGCCCTTGAGAGCCCGCGCGGCT72

QY1668AGTTACCATCACACCCCGGAGGCGCGCAG1698

Db73AGTTACCATCACACCCCGGAGGCGCGCAG103

RESULT 4

US-09-880-107-2100

Sequence 2100, Application US/09880107

Patent No. US20020142981A1

GENERAL INFORMATION:

APPLICANT: Horne, Darci T.

APPLICANT: Vockley, Joseph G.

APPLICANT: Scherf, Uwe

APPLICANT: Gene Logic, Inc.

TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer

FILE REFERENCE: 44921-5028-WO

CURRENT APPLICATION NUMBER: US/09/880,107

CURRENT FILING DATE: 2001-06-14

PRIOR APPLICATION NUMBER: US 60/211,379

PRIOR FILING DATE: 2000-06-14

PRIOR APPLICATION NUMBER: US 60/237,054

PRIOR FILING DATE: 2000-10-02

NUMBER OF SEQ ID NOS: 3950

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2100

LENGTH: 1481

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: Genbank Accession No. US20020142981A1 J03827

US-09-880-107-2100

Query Match5.2%; Score 89; DB 9; Length 1481;

Best Local Similarity100.0%; Pred. No. 3e-17;

Matches89; Conservative0; Mismatches0; Indels0; Gaps0;

QY1610CGGGAGCGGAGCGGACCCCGAGAGCCCTTGAGAGCCCGCGCGGCTAG1669

Db2CGGGAGCGGAGCGGACCCCGAGAGCCCTTGAGAGCCCGCGCGGCTAG61

QY1670TTACCATCACACCCCGGAGGCGCGCAG1698

Db62TTACCATCACACCCCGGAGGCGCGCAG90

RESULT 5

US-10-023-969-2

Sequence 2, Application US/10023969

Publication No. US2003009598A1

GENERAL INFORMATION:

APPLICANT: Geron Corporation

APPLICANT: Irving, John

APPLICANT: Lebkowski, Jane

TITLE OF INVENTION: Chimeric Cytolytic Viruses for Cancer Treatment

FILE REFERENCE: 084,002

CURRENT APPLICATION NUMBER: US/10/023,969

CURRENT FILING DATE: 2001-12-17

PRIOR APPLICATION NUMBER: 60/256,418

PRIOR FILING DATE: 2000-12-18

NUMBER OF SEQ ID NOS: 4

SOFTWARE: PatentIn version 3.1

SEQ ID NO 2

LENGTH: 1481

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (127)..(1080)

US-10-023-969-2

Query Match5.3%; Score 89.4; DB 14; Length 395;

Best Local Similarity98.9%; Pred. No. 8.7e-18;

Matches90; Conservative0; Mismatches1; Indels0; Gaps0;

```
; NAME/KEY: misc_feature
; OTHER INFORMATION: Template ID: 1000222.1
US-09-971-392-82

Query Match      5.2%; Score 89; DB 14; Length 1481;
Best Local Similarity 100.0%; Pred. No. 3e-17; Indels 0; Gaps 0;
Matches 89; Conservative 0; Mismatches 0;

QY 1610 CGGAGCGGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCAACCGCGCGCGCTAG 1669
Db 2 CGGAGCGGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCAACCGCGCGCGCTAG 61

QY 1670 TTACCATCACACCCCGGAGAGCGCGCAG 1698
Db 62 TTACCATCACACCCCGGAGAGCGCGCAG 90

RESULT 6
US-09-764-891-2685
; Sequence 2685, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2685
; LENGTH: 179
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (117)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-891-2685

Query Match      5.2%; Score 88.6; DB 10; Length 179;
Best Local Similarity 95.8%; Pred. No. 9.1e-18; Indels 0; Gaps 0;
Matches 91; Conservative 0; Mismatches 4;

QY 1604 CGGTAGCGGAGCGGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCAACCGCGCGCG 1663
Db 18 CGTGGCGGAGCGGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCAACCGCGCGCG 77

QY 1664 GCTAGTTACCATCACACCCCGGAGAGCGCGCAG 1698
Db 78 GCTAGTTACCATCACACCCCGGAGAGCGCGTAG 112

RESULT 7
US-09-971-392-82
; Sequence 82, Application US/09971392
; Publication No. US20030134283A1
; GENERAL INFORMATION:
; APPLICANT: Peterson, David P.
; APPLICANT: Pearson, Cecelia I.
; APPLICANT: Cocks, Benjamin G.
; TITLE OF INVENTION: GENES REGULATED IN DENDRITIC CELL DIFFERENTIATION
; FILE REFERENCE: PA-0029 US
; CURRENT APPLICATION NUMBER: US/09/971,392
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: 60/237,652
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 260
; SOFTWARE: PERL Program
; SEQ ID NO 82
; LENGTH: 1520
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: misc_feature
; OTHER INFORMATION: Template ID: 1000222.1
US-09-971-392-82

Query Match      5.1%; Score 87.4; DB 10; Length 1520;
Best Local Similarity 98.0%; Pred. No. 1e-16; Indels 1; Gaps 1;
Matches 99; Conservative 0; Mismatches 1;

QY 1599 TCATCGGTAGCGGAGCGGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCAACCGCGCG 1658
Db 5 TAGATCGGTAGCGGAGCGGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCAACCGCGCG 64

QY 1659 CGCGGCGCTAGTTACCATCACACCCCGG-GAGGAGCGCGCAG 1698
Db 65 CGCGGCGCTAGTTACCATCACACCCCGGTGAGGAGCGCGCAG 105

RESULT 8
US-09-738-973-400
; Sequence 400, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raedoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0.
; SEQ ID NO 400
; LENGTH: 410
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-738-973-400

Query Match      5.0%; Score 85.2; DB 9; Length 410;
Best Local Similarity 96.7%; Pred. No. 2.1e-16; Indels 0; Gaps 0;
Matches 87; Conservative 0; Mismatches 3;

QY 1609 GCAGGAGCGGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCAACCGCGCGCGCTA 1668
Db 2 GCAGGAGCGGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCAACCGCGCGCGCTA 61

QY 1669 GTTACCATCACACCCCGGAGAGCGCGCAG 1698
Db 62 GTTACCATCACACCCCGGAGAGCGCGCAG 91

RESULT 9
US-09-854-133-400
; Sequence 400, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raedoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
```

```

; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 400
; LENGTH: 410
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-854-133-400

Query Match      5.0%; Score 85.2; DB 9; Length 410;
Best Local Similarity 96.7%; Pred. No. 2.1e-16;
Matches 87; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1609 GCGGAGCGGAGGGAGCCGCCAGAGAGCCTTGACAGCCCCACCGCGCGCGGCCTA 1668
DB 2 GCACGAGGGAGAGCGGAGCCCGCCAGAGAGCCTTGACAGCCCCACCGCGCGCGGCCTA 61

QY 1669 GTTACCATCACACCCCAGGAGGAGCGGCAG 1698
DB 62 GTTACCATCACACCCCAGGAGGAGCGGCAG 91

RESULT 10
US-10-144-649A-400
; Sequence 400, Application US/10144649A
; Publication No. US20030118599A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Algate, Paul A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C11
; CURRENT APPLICATION NUMBER: US/10/144,649A
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 749
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 400
; LENGTH: 410
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-144-649A-400

Query Match      5.0%; Score 85.2; DB 14; Length 410;
Best Local Similarity 96.7%; Pred. No. 2.1e-16;
Matches 87; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1609 GCGGAGCGGAGAGCGGAGCCCGCAGAGAGCCTTGACAGCCCCACCGCGCGCGGCCTA 1668
DB 2 GCACGAGGGAGAGCGGAGCCCGCCAGAGAGCCTTGACAGCCCCACCGCGCGCGGCCTA 61

QY 1669 GTTACCATCACACCCCAGGAGGAGCGGCAG 1698
DB 62 GTTACCATCACACCCCAGGAGGAGCGGCAG 91

RESULT 11
US-10-424-599-138143
; Sequence 138143, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 138143
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_95751C.1
US-10-424-599-138143

Query Match      4.9%; Score 84; DB 12; Length 492;
Best Local Similarity 100.0%; Pred. No. 5.8e-16;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1615 GCGGAGAGGGAGCCCGCAGAGAGCCTTGACAGCCCCACCGCGCGCGGCCTAGTTACC 1674
DB 1 GCGGAGAGGGAGCCCGCAGAGAGCCTTGACAGCCCCACCGCGCGCGGCCTAGTTACC 60

QY 1675 ATCACACCCCGGAGGAGCGGCAG 1698
DB 61 ATCACACCCCGGAGGAGCGGCAG 84

RESULT 12
US-09-735-705-215
; Sequence 215, Application US/09735705
; Patent No. US20020052329A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fang, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C14
; CURRENT APPLICATION NUMBER: US/09/735,705
; CURRENT FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 215
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: unsure
; LOCATION: (17)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (20)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (60)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (61)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (365)
; OTHER INFORMATION: n=A,T,C or G
US-09-735-705-215

Query Match      4.6%; Score 78.8; DB 9; Length 381;
Best Local Similarity 93.0%; Pred. No. 2.3e-14;
Matches 80; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1613 GAGCGGAGAGCGGAGCCCGCAGAGAGCCTTGACAGCCCCACCGCGCGGCCTAGTTA 1674

```

Db 1 GAGCGAGAGCGGACCGTGTAGAGCCCTGAGCAGCCCCACCGCGCCGCGCTAGTTN 60
QY 1673 CCATCACACCCCGGAGGAGCGCGAG 1698
Db 61 NCATCACACCCCGGAGGAGCGCGAG 86

RESULT 13

US-09-850-716A-215
; Sequence 215, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 215
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (17)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (20)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (60)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (61)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (365)
; OTHER INFORMATION: n=A,T,C or G
US-09-850-716A-215

Query Match 4.6%; Score 78.8; DB 9; Length 381;
Best Local Similarity 93.0%; Pred. No. 2.3e-14;
Matches 80; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1613 GAGCGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCCACCGCGCCGCGCTAGTTA 1672
Db 1 GAGCGAGAGCGGACCGTGTAGAGCCCTGAGCAGCCCCACCGCGCCGCGCTAGTTN 60
QY 1673 CCATCACACCCCGGAGGAGCGCGAG 1698
Db 61 NCATCACACCCCGGAGGAGCGCGAG 86

RESULT 14

US-09-897-778-215
; Sequence 215, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Mainerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 215
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 17, 20, 60, 61, 365
; OTHER INFORMATION: n = A,T,C or G
US-09-897-778-215

Query Match 4.6%; Score 78.8; DB 9; Length 381;
Best Local Similarity 93.0%; Pred. No. 2.3e-14;
Matches 80; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1613 GAGCGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCCACCGCGCGCGCTAGTTA 1672
Db 1 GAGCGAGAGCGGACCGTGTAGAGCCCTGAGCAGCCCCACCGCGCGCGCTAGTTN 60
QY 1673 CCATCACACCCCGGAGGAGCGCGAG 1698
Db 61 NCATCACACCCCGGAGGAGCGCGAG 86

RESULT 15

US-09-466-396A-215
; Sequence 215, Application US/09466396A
; Publication No. US20030119763A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.455C4
; CURRENT APPLICATION NUMBER: US/09/466,396A
; CURRENT FILING DATE: 1999-12-17
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 215
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (17)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (20)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (60)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (61)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (365)
; OTHER INFORMATION: n=A,T,C or G
US-09-466-396A-215

Query Match 4.6%; Score 78.8; DB 10; Length 381;
Best Local Similarity 93.0%; Pred. No. 2.3e-14;
Matches 80; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1613 GAGCGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCCACCGCGCGCGCTAGTTA 1672
Db 1 GAGCGAGAGCGGACCGTGTAGAGCCCTGAGCAGCCCCACCGCGCGCGCTAGTTN 60
QY 1673 CCATCACACCCCGGAGGAGCGCGAG 1698

Db 61 NCATCACCCCCGGGAGGAGCCGAG 86

Search completed: March 28, 2004, 19:01:23
Job time : 654 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2004, 13:35:55 ; Search time 4640 Seconds
(without alignments)
10928.013 Million cell updates/sec

Title: US-09-869-508-1_COPY_453_2150

Perfect score: 1698

Sequence: 1 agaacattatggtcgccaa.....caccocggaggagccgcag 1698

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estnu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rpd.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gsl.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	519.4	30.6	532	9	AI659243
C 2	489.4	28.8	503	12	BI830189
C 3	451.6	26.6	565	10	BF111618
C 4	451.6	26.6	608	13	BX114760

C	5	451.6	26.6	745	14	CB850550
C	6	436	25.7	509	9	AA463628
C	7	425.4	25.1	438	9	AI650982
C	8	418.4	24.6	431	9	AI268811
C	9	410.4	24.2	428	9	AI467820
C	10	405.8	23.9	428	9	AI208001
C	11	399.4	23.5	412	14	N67996
C	12	354.2	20.9	371	10	AM448961
C	13	354.2	20.9	378	10	BE585954
C	14	352.8	20.8	866	10	BF570594
C	15	332	19.6	434	12	BI828640
C	16	327.2	19.3	543	29	CE384034
C	17	217.6	12.8	238	9	AI640509
C	18	190	11.2	803	28	AO746664
C	19	187.4	11.0	333	28	BZ867318
C	20	186.8	11.0	1016	11	AK007113
C	21	186.2	11.0	977	13	BY707189
C	22	164	9.7	703	14	CB467717
C	23	140.6	8.3	279	12	BI828773
C	24	135	8.0	600	10	BB615022
C	25	133.4	7.9	217	12	BI550827
C	26	128	7.5	913	13	EX422060
C	27	124.8	7.3	1000	13	EX450069
C	28	116	6.8	508	9	AUI26092
C	29	116	6.8	1136	12	BM459919
C	30	114.4	6.7	610	12	BI830400
C	31	112	6.6	403	12	BM853312
C	32	112	6.6	476	12	BM742560
C	33	111.4	6.6	1219	12	BM477621
C	34	110	6.5	643	10	AW163133
C	35	108	6.4	925	13	BQ421211
C	36	108	6.4	960	12	EG387948
C	37	108	6.4	1228	10	BS436330
C	38	106	6.2	950	13	BQ214006
C	39	106	6.2	1403	10	BF339614
C	40	105.4	6.2	715	12	BG773027
C	41	104.4	6.1	762	9	AI878912
C	42	103	6.1	165	12	BM847680
C	43	103	6.1	166	12	BM739099
C	44	103	6.1	358	12	BM851331
C	45	103	6.1	360	12	BM850156

ALIGNMENTS

RESULT 1
AI659243/c
LOCUS tui8b06.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:2251379 3',
DEFINITION mRNA sequence.
ACCESSION AI659243
VERSION AI659243.1 GI:4762813
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 532)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs@mail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

532 bp mRNA linear EST 10-MAY-1999

AI659243 tui8b06.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:2251379 3',
DEFINITION mRNA sequence.

ACCESSION AI659243
VERSION AI659243.1 GI:4762813

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Seq primer: -40UP from Gibco
High quality sequence stop: 463.
Location/Qualifiers
1. 532

FEATURES

source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2251379"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="NCI CGAP Pr28"
/notes="Organ: prostate; Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP Pr22 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clones 985608-986759, 1101192-1101959, and 1217928-1220615). Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 30.6%; Score 519.4; DB 9; Length 532;
Best Local Similarity 99.8%; Pred. No. 5.9e-137;
Matches 531; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 245 TCCTACTTCCATTCCTTGCCTCTGCAACCAATCTCCATCTCGGAGCCATTTTAA 304
DB 532 TCCTACTTCCATTCCTTGCCTCTGCAACCAATCTCCATCTCGGAGCCATTTTAA 473
QY 305 AAGATCCCTCCCTACTATGACTCTAAATGTCTTCTCACTTCCCTCAGGATA 364
DB 472 AAGATCCCTCCCTACTATGACTCTAAATGTCTTCTCACTTCCCTCAGGATA 413
QY 365 TATTTCAATTAATATACCTAAGTACTGCCACCTCTGCAACCAATGTCAATTCGA 424
DB 412 TATTTCAATTAATATACCTAAGTACTGCCACCTCTGCAACCAATGTCAATTCGA 353
QY 425 GTCTTACTGACTTACTGACTTCCCGAGATCTCACTTCTTCCCTGTACCTGT 484
DB 352 GTCTTACTGACTTACTGACTTCCCGAGATCTCACTTCTTCCCTGTACCTGT 293
QY 485 TGGCGGAAAGTCAGCGCTCCACCTTCTCCCTCTCCACTCCCAAAATCTCTGTGT 544
DB 292 TGGCGGAAAGTCAGCGCTCCACCTTCTCCCTCTCCACTCCCAAAATCTCTGTGT 233
QY 545 TTGCAGCTCTGGAGTATTACCGTGTGGCTGTGTTAAATTTCTGCTCCATCAGAGGCA 604
DB 232 TTGCAGCTCTGGAGTATTACCGTGTGGCTGTGTTAAATTTCTGCTCCATCAGAGGCA 173
QY 605 GAACTGACTCGGACTATTCCATCCCGCCGATAGTAGCGCTTAAAGAGACGA 564
DB 172 GAACTGACTCGGACTATTCCATCCCGCCGATAGTAGCGCTTAAAGAGACGA 113
QY 665 AGAAGTGGGTGGAGGACTTCACTAGTACATCAGTGGGAGCTCAATTTTATCGTTGTG 724
DB 112 AGAAGTGGGT - GAGAGACTTCACTAGTACATCAGTGGGAGCTCAATTTTATCGTTGTG 54
QY 725 AAACGTGGATAGTAATCCCTCTATCAGTGGCTGTGTCAGGATATAAGTGAA 777
DB 53 AAACGTGGATAGTAATCCCTCTATCAGTGGCTGTGTCAGGATATAAGTGAA 1

RESULT 2

BI830189/c
LOCUS 503 bp mRNA linear EST 04-OCT-2001
DEFINITION 603072855F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5164981 5', mRNA sequence.
ACCESSION BI830189
VERSION BI830189.1 GI:15941739
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 503)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-femail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: InCyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM11409 row: m column: 14
High quality sequence stop: 499.
Location/Qualifiers
1. 503

REFERENCE

AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES

source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5164981"
/tissue_type="medulla"
/lab_host="DH10B"
/clone_lib="NIH_MGC_119"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 28.8%; Score 489.4; DB 12; Length 503;
Best Local Similarity 99.6%; Pred. No. 2.2e-128;
Matches 501; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 698 GTGCAGCCTCAATTTATCGTTTGGAAACGTTAGTAGTAATCCCTCTATCAGTGGCT 757
DB 502 GTGCAGCCTCAATTTATCGTTTGGAAACGTTAGTAGTAATCCCTCTATCAGTGGCT 443
QY 758 GTTCAGAGATAAGTCAAAAACAAACAGGCTAGCTTCTCAATAAATGTCAGTTGAA 817
DB 442 GTTCAGAGATAAGTG - AAAAATAAACAGGCTAGCTTCTCAATAAATGTCAGTTGAA 384
QY 818 TTAATCTGATTTGTGTCAGTAGAGAAAAGATGTGAATCTTGGAAAGGAAGACACATTT 877
DB 383 TTAATCTGATTTGTGTCAGTAGAGAAAAGATGTGAATCTTGGAAAGGAAGACACATTT 324
QY 878 TTTTAAATATATGCTGGTAAACGGATCAGAGGAGGTCCTCCATGAGACACACCTCG 937
DB 323 TTTTAAATATATGCTGGTAAACGGATCAGAGGAGGTCCTCCATGAGACACACCTCG 264
QY 938 CCTTAAACATGCTGAACCCGGCTGCCATAGCTTCCCTGCTCCCTCAAGTGACTGCTC 997
DB 263 CCTTAAACATGCTGAACCCGGCTGCCATAGCTTCCCTGCTCCCTCAAGTGACTGCTC 204
QY 998 CGACAAAAGGTCAGCTCTTCAAAACGCATACGTTTAAAGGCAATTCAGAAACCCCTCGGCT 1057
DB 203 CGACAAAAGGTCAGCTCTTCAAAACGCATACGTTTAAAGGCAATTCAGAAACCCCTCGGCT 144
QY 1058 GTGCGCGGACTACACGCCCATTTAAAGAAAAGACGACTCTATGCGCCGCGTAATGTTCTCA 1117
DB 143 GTGCGCGGACTACACGCCCATTTAAAGAAAAGACGACTCTATGCGCCGCGTAATGTTCTCA 84
QY 1118 GATCACAGGACCGTATTATGGAGCTGGAGGGAGGAGCGCTTTTCTTTCAGCGGGGGCTA 1177
DB 83 GATCACAGGACCGTATTATGGAGCTGGAGGGAGGAGCGCTTTTCTTTCAGCGGGGGCTA 24

```
QY 1178 AGGGGTCTTCAGGCCCTTCCA 1200
|||||
DB 23 AGGGGTCTTCAGGCCCTTCCA 1

RESULT 3
BF111618/c
LOCUS BF111618 565 bp mRNA linear EST 20-OCT-2000
DEFINITION IMAGE:3523180.3, mRNA sequence.
ACCESSION BF111618
VERSION BF111618.1 GI:10941308
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 565)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 463.
FEATURES
Location/Qualifiers
1..565
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3523180"
/lab_host="DH10B"
/clone_lib="Soares NSF P8 9W OT PA P S1"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and clones: Soares NB2HP pool 1:
309384-310919, 323208-325895 Soares NB2HP pool 1:
145032-147335, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares NB2HP8-9W pool 1:
758280-760583, 772104-774407 Soares NBHPA pool 1:
304776-306311, 320136-322823, 326280-326663 Soares NBHOT
pool 1: 723720-726407, 739080-740999 Subtraction by Bento
Soares and M. Fatima Bonaldo."
ORIGIN
Query Match 26.6%; Score 451.6; DB 10; Length 565;
Best Local Similarity 98.9%; Pred. No. 1.5e-117;
Matches 465; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 353 CCCCTCAGGATATATTCCAATTAATACTAAGTGAAGTCCACCTCTGCAACCCAA 412
DB 472 CACTTGAGGATATATTCCAATTAATACTAAGTGAAGTCCACCTCTGCAACCCAA 413
QY 413 TGTACATTCAGTCTTACTCACTTGAAGTCTGATTCGATTCCTCCGAGATCTCACTCTTCTC 472
DB 412 TGTACATTCAGTCTTACTGAATCTGATCTGATTCGATTCCTCCGAGATCTCACTCTTCTC 353
QY 473 GCCTGTACCTGTGCGCGAAAGTCAGCCCTCCACCTTCTCCCTGCTTCACTCCCAAAA 532
DB 352 GCCTGTACCTGTGCGCGAAAGTCAGCCCTCCACCTTCTCCCTGCTTCACTCCCAAAA 293
QY 533 TACTTGTGGTTCAGCTCTGAGATATTTACCGTGTGGCTGTAAATTTCTGCTC 592
|||||
```

```
Db 292 TACTTCGTGTTTTTGCAGCTCTGGAGTATTTACCGTGTGGCTGTAAATTTCTGCTC 233
QY 593 CATCAGAGGCGAGAACTGACTCGGGAATTTTCCATCCCGAGCGATAGTAGCGCTTA 652
DB 232 CATCAGAGGCGAGAACTGACTCGGGAATTTTCCATCCCGAGCGATAGTAGCGCTTA 173
QY 653 AAAAGAACGGAAGAGGTGGTGGGAGGACTTCAGTAAACATCAGGTGGCAGCCCTCAATT 712
DB 172 AAAAGAACGGAAGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 114
QY 713 TTATCGTTTGTGAACGCTGAGTAGTAATCCCTCTATCAGCTGGCTGTGTGCGAGATAAAG 772
DB 113 TTATCGTTTGTGAACGCTGAGTAGTAATCCCTCTATCAGCTGGCTGTGTGCGAGATAAAG 54
QY 773 TGAATAAACAAACAGGCTAGCTTCTCAATAAATGTGAGTTGAATTTAA 822
DB 53 TGAATAAACAAACAGGCTAGCTTCTCAATAAATGTGAGTTGAATTTAA 4

RESULT 4
LOCUS BX114760 608 bp mRNA linear EST 07-FEB-2003
DEFINITION BX114760 Soares NHMPu S1 Homo sapiens cDNA clone IMAGE:9998H112004 ;
IMAGE:811834, mRNA sequence.
ACCESSION BX114760
VERSION BX114760.1 GI:27879797
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 608)
Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
Radelof, U., Schneider, D. and Korn, B.
Human Unigeneset - RZPD3
Unpublished (2003)
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD LIB; I.M.A.G.E. cDNA Clone Collection;
Human Unigeneset - RZPD3 (RZPD LIB No.972)
http://www.rzpd.de/CloneCards/cgi-
bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTCACAGGAAACAGCTATGAC.
FEATURES
Location/Qualifiers
1..608
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:9998H112004 ; IMAGE:811834"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/clone_lib="Soares NHMPu S1"
/note="Organ: mixed (see below); Vector: pT73D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not 1;
Site 2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBHM, pregnant uterus
NBHPu, and fetal heart NBH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."
```

```
ORIGIN
Query Match      26.6%; Score 451.6; DB 13; Length 608;
Best Local Similarity 98.9%; Pred. No. 1.6e-117;
Matches 465; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

353 CCCTCAGGATATATTTCCAAATTAATATACCTAACTAGTACCTGCCACCTCTGCAACCCAA 412
Db      128 CACTTGAGGATATATTTCCAAATTAATATACCTAACTAGTACCTGCCACCTCTGCAACCCAA 187
Qy      413 TGTCAATTCGAGTCTTACTGAACACTACTGACATTCATTCCTCCGAGATCTCACTCTTCTC 472
Db      188 TGTCAATTCGAGTCTTACTGAACACTACTGACATTCATTCCTCCGAGATCTCACTCTTCTC 247
Qy      473 GCCTGTACCTCTGCGGGAAGTCAAGCTCCACCTCTCCCTGCTTCCATCCCAAA 532
Db      248 GCCTGTACCTCTGCGGGAAGTCAAGCTCCACCTCTCCCTGCTTCCATCCCAAA 307
Qy      533 TACTTCGTGTTTGGAGCTCTGAGATTAATACCGTGTGGCTGTGTTAAATTTCTGCCCTC 592
Db      308 TACTTCGTGTTTGGAGCTCTGAGATTAATACCGTGTGGCTGTGTTAAATTTCTGCCCTC 367
Qy      593 CATCAGAAGCGAAGAACTGACTCGGAACATTAATCCATCCCGAGCGATAGAGCTTTA 652
Db      368 CATCAGAAGCGAAGAACTGACTCGGAACATTAATCCATCCCGAGCGATAGAGCTTTA 427
Qy      653 AAAAAGAACGGAAGAGTGGTGGAGGACTTCAGTAACATCAGTGGCAGCTCAATT 712
Db      428 AAAAAGAACGGAAGAGTGGGT-GGAGGACTTCAGTAACATCAGTGGCAGCTCAATT 486
Qy      713 TTATCGTTTGTGAACGTTGATAGTAATCCCTCTATCAGTGGCTGTTCGAGGAATAAAG 772
Db      487 TTATCGTTTGTGAACGTTGATAGTAATCCCTCTATCAGTGGCTGTTCGAGGAATAAAG 546
Qy      773 TGAATAAACAAAAACAGGCTAGCTGTGTTCAATAAATGTGAGTTCAATTA 822
Db      547 TGAATAAACAAAAACAGGCTAGCTGTGTTCAATAAATGTGAGTTCAATTA 596

RESULT 5
CB850550/c
LOCUS
DEFINITION
  UI-CP-EN1-adc-h-05-0-UI.s1 UI-CP-EN1 Homo sapiens cDNA clone
  UI-CP-EN1-adc-h-05-0-UI.3', mRNA sequence.
ACCESSION
  CB850550
  CB850550.1 GI:30045317
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 745)
  Bonaldo,M.F., Lennon,G. and Soares,M.B.
  Normalization and subtraction: two approaches to facilitate gene
  discovery
  Genome Res. 6 (9), 791-806 (1996)
  97044477
  8889548
  Contact: McCray, PB
  McCray Lab
  University of Iowa
  2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
  Tel: 319 356 4866
  Fax: 319 356 7171
  Email: paul-mccray@uiowa.edu
  Tissue Procurement: Dr. M. J. Welsh, University of Iowa
  cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
  cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: Researchers may obtain clones from Research
  Genetics (www.resgen.com) or from Open Biosystems
  (www.openbiosystems.com).
  The following repetitive elements were found in this cDNA
```

```
sequence: 1-22, >AT_rich#Low_complexity 542-620, >LINE2 (matched
compliment)
Seq primer: M13 FORWARD
POLYA=Yes.
FEATURES
  source
  1..745
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="UI-CP-EN1-adc-h-05-0-UI"
  /tissue_type="Primary Lung Cystic Fibrosis Epithelial
  Cells"
  /dev_stage="Adult"
  /lab_host="DH10B (Life Technologies) (TI phage resistant)"
  /clone_lib="UI-CP-EN1"
  /note="Torgan; Lung; Vector: pT73-Pac (Pharmacia) with a
  modified polylinker; Site 1: Sc01; Site 2: Not 1;
  UI-CP-EN1 is a normalized cDNA library containing the
  following tissue(s): Primary Lung Cystic Fibrosis
  Epithelial Cells. The library was constructed according to
  Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
  1996. First strand cDNA synthesis was primed with an
  oligo-dT primer containing a Not I site. Double stranded
  cDNA was ligated to an EcoR I adaptor, digested with Not
  I, and cloned directionally into pT73-Pac vector. The
  oligonucleotide used to prime the synthesis of
  first-strand cDNA contains a library tag sequence that is
  located between the Not I site and the (dT)18 tail. The
  sequence tag for this library is CTGCTCAGGT.
  TAG_SEQ=None found"
ORIGIN
Query Match      26.6%; Score 451.6; DB 14; Length 745;
Best Local Similarity 98.9%; Pred. No. 1.7e-117;
Matches 465; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy      353 CCCCTCAGGATATATTTCCAAATTAATATACCTAACTAGTACCTGCCACCTCTGCAACCCAA 412
Db      484 CACTTGAGGATATATTTCCAAATTAATATACCTAACTAGTACCTGCCACCTCTGCAACCCAA 425
Qy      413 TGTCAATTCGAGTCTTACTGAACACTACTGACTCCATTCCTCCGAGATCTCACTCTTCTC 472
Db      424 TGTCAATTCGAGTCTTACTGAACACTACTGACTCCATTCCTCCGAGATCTCACTCTTCTC 365
Qy      473 GCCTGTACCTCTGCGGGAAGTCAAGCTCCACCTCTCCCTGCTTCCATCCCAAA 532
Db      364 GCCTGTACCTCTGCGGGAAGTCAAGCTCCACCTCTCCCTGCTTCCATCCCAAA 305
Qy      533 TACTTCGTGTTTGGAGCTCTGAGATTAATACCGTGTGGCTGTGTTAAATTTCTGCCCTC 592
Db      304 TACTTCGTGTTTGGAGCTCTGAGATTAATACCGTGTGGCTGTGTTAAATTTCTGCCCTC 245
Qy      593 CATCAGAAGCGAAGAACTGACTCGGAACATTAATCCATCCCGAGCGATAGAGCTTTA 652
Db      244 CATCAGAAGCGAAGAACTGACTCGGAACATTAATCCATCCCGAGCGATAGAGCTTTA 185
Qy      653 AAAAAGAACGGAAGAGTGGTGGAGGACTTCAGTAACATCAGTGGCAGCTCAATT 712
Db      184 AAAAAGAACGGAAGAGTGGGT-GGAGGACTTCAGTAACATCAGTGGCAGCTCAATT 126
Qy      713 TTATCGTTTGTGAACGTTGATAGTAATCCCTCTATCAGTGGCTGTTCGAGGAATAAAG 772
Db      125 TTATCGTTTGTGAACGTTGATAGTAATCCCTCTATCAGTGGCTGTTCGAGGAATAAAG 66
Qy      773 TGAATAAACAAAAACAGGCTAGCTGTGTTCAATAAATGTGAGTTCAATTA 822
Db      65 TGAATAAACAAAAACAGGCTAGCTGTGTTCAATAAATGTGAGTTCAATTA 16

RESULT 6
AA463628/c
LOCUS
DEFINITION
  AA463628 509 bp mRNA linear EST 10-JUN-1997
  zx98906.s1 Soares_NHMPu_S1 Homo sapiens cDNA clone IMAGE:811834
```

3', mRNA sequence.
AA463628 1 GI:2188512
EST.
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisels, G., Jost, S.,
Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J.,
Moore, B., Schellenberg, K., Steptoe, W., Tan, F., Theising, J.,
White, Y., Wyllie, T., Waterston, R. and Wilson, R.
WashU-Merck EST Project 1997
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -4ln13 fwd. ET from Amersham
High quality sequence stop: 486.
FEATURES
Location/Qualifiers
1..509
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:6042745"
/db_xref="taxon:9606"
/clone="IMAGE:811834"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/clone_lib="Soares NHMPU S1"
/note="Organ: mixed (see below); Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not 1;
Site 2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBHM, pregnant uterus
NbHPU, and fetal heart NbH19W) were mixed, and ss circles
was made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."
ORIGIN
Query Match 25.7%; Score 436; DB 9; Length 509;
Best Local Similarity 98.5%; Pred. No. 4.3e-113;
Matches 461; Conservative 0; Mismatches 5; Indels 2; Gaps 2;
QY 353 CCCCTCAGGATATATTTCCCAATTAATACCTAAGTGACTGCCACCTCTGCAACCCAA 412
DB 467 CACTGTGAGATATATTTCCCAATTAATACCTAAGTGACTGCCACCTCTGCAACCCAA 408
QY 413 TGTACATTCGAGTCTTACTGAACTACTTGACTGCAATTCGCGAGATCTCACTCTTCTC 472
DB 407 TGTACATTCGAGTCTTACTGAACTACTTGACTGCAATTCGCGAGATCTCACTCTTCTC 348
QY 473 GCCTGTACCTGTGCGGGGAAAGTCAGCCCTTCCACCTTCTCCCTGCTTCCACTCCC-AAA 531
DB 347 GCCTGTACCTGTGCGGGGAAAGTCAGCCCTTCCACCTTCTCCCTGCTTCCACTCCCAGAA 288
QY 532 ATACTTGTGTTTGTGAGCTCTGAGTATTACCTGTTGGCTGTTAAATTTCTGCCT 591
DB 287 ATACTTGTGTTTGTGAGCTCTGAGTATTACCTGTTGGCTGTTAAATTTCTGCCT 228
QY 592 CCATCAGAGCGAAGAACTGACTCGCGAAGTATTCCATCCCGACCGATAGTAGAGCGCTT 651
DB 227 CCATCAGAGCGAAGAACTGACTCGCGAAGTATTCCATCCCGACCGATAGTAGAGCGCTT 168

652 AAAAAAGAACGAAAGAGGTGGTGGAGGACTTCAGTAACATCAGGTGGAGGCTCAAT 711
167 AAAAAAGAACGAAAGAGGTGGT- GGAGGACTTCAGTAACATCAGGTGGAGGCTCAAT 109
712 TTTATCGTTTGTGAACGTTGATAGTAACTCCCTCTATCAGCTGGCTGTTGAGGGAATAA 771
108 TTTATCGTTTGTGAACGTTGATAGTAACTCCCTCTATCAGCTGGCTGTTGAGGGAATAA 49
772 GTGAAAAACAAACACAGGCTAGCTTGTTCATATAATGTGAGTTGAATT 819
48 GTGAAAAACAAACACAGGCTAGCTTGTTCATATAATGTGAGTTGAATT 1
RESULT 7
LOCUS
AI650982 438 bp mRNA linear EST 16-DEC-1999
wa96d06.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2304011 3',
mRNA sequence.
ACCESSION
AI650982
VERSION
AI650982.1 GI:4734961
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. (bases 1 to 438)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 552 Std Error: 0.00
Seq primer: -40UP from Gibco.
FEATURES
Location/Qualifiers
1..438
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2304011"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/clone_lib="NCI CGAP GC6"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not 1; Site 2: Eco RI; Plasmid DNA
from the normalized library NCI_CGAP_GC4 was prepared, and
ss circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneIDs:
1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo."
ORIGIN
Query Match 25.1%; Score 425.4; DB 9; Length 438;
Best Local Similarity 99.8%; Pred. No. 4.4e-110;
Matches 426; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 858 TTGGAAGGAGACACATTTTAAATATATATGCTGTTAAACCGATCAGAAGCAGGT 917
DB 3 TTGGAAGGAGACACATTTTAAATATATATGCTGTTAAACCGATCAGAAGCAGGT 62
QY 918 CCCCTGAGGACACCCCTGCGCTAAACATGCTGAACCCGGGCTGCCATAGCTCGTGG 977

a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from five normalized libraries were mixed, and es circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 5 libraries. The pools consisted of the following libraries and clones: Soares NbHSF pool 1: 309384-310919, 323208-325895 Soares Nb2HP pool 1: 145032-147335, 147720-148103, 148872-149255, 15002 - 150407, 151176-152327 Soares Nb2HF8-9W pool 1: 758280-760583, 772104-774407 Soares NbHFA pool 1: 304776-306311, 320136-322823, 326280-326663 Soares NbHOT pool 1: 723720-726407, 739080-740999 Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 24.2%; Score 410.4; DB 9; Length 428;
Best Local Similarity 99.5%; Pred. No. 8.5e-106;
Matches 422; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 399 CCTGCAACCAATGTCACATTCGAGCTTACTGCACTACTTGAAGTCTTCCCGAGA 458
Db CCTGCAACCAATGTCACATTCGAGCTTACTGCACTACTTGAAGTCTTCCCGAGA 369

QY 459 TCTCACTCTCTCGCTGTACCTGTGCGGGAAGTCAAGCTTCCACCTTCCCTGC 518
Db TCTCACTCTCTCGCTGTACCTGTGCGGGAAGTCAAGCTTCCACCTTCCCTGC 309

QY 519 TTCACTCTCCCAATACCTGCTGTTTTCGAGCTCTGAGTATTTACCGTCTTGGCTGT 578
Db TTCACTCTCCCAATACCTGCTGTTTTCGAGCTCTGAGTATTTACCGTCTTGGCTGT 249

QY 579 TAAATTTCTGCTCCATCAGAGGAGCAAACTGACTCGCAACTATTCCATCCCGACG 638
Db TAAATTTCTGCTCCATCAGAGGAGCAAACTGACTCGCAACTATTCCATCCCGACG 189

QY 639 ATAGTAGAGCTTAAAGAACGCAAGAGTGGTGGGAGGACTTCAGTAAATCAGG 698
Db ATAGTAGAGCTTAAAGAACGCAAGAGTGGTGGGAGGACTTCAGTAAATCAGG 130

QY 699 TGGCAGCTCAATTTATCGTTTGTGAAAGTGTAGTATTCCTCTATCAGCTGGCTG 758
Db TGGCAGCTCAATTTATCGTTTGTGAAAGTGTAGTATTCCTCTATCAGCTGGCTG 70

QY 759 TTGCAGGAATAAGTGAAAAACAAACAGAGCTAGCTTTGTTCAATAATGTGAGTTGAAT 818
Db TTGCAGGAATAAGTGAAAAACAAACAGAGCTAGCTTTGTTCAATAATGTGAGTTGAAT 10

QY 819 TAAA 822
Db 9 TAAA 6

RESULT 10
A1208001/c
LOCUS
DEFINITION 428 bp mRNA linear EST 30-NOV-1998
3', mRNA sequence.
ACCESSION A1208001
VERSION A1208001.1 GI:3769943
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 428)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center.
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/brp/image/image.html
Insert Length: 603 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 418.
Location/Qualifiers
source 1. 428
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1839586"
/sex="male"
/lab_host="DH10B"
/clone_lib="Soares testis NHT"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo (dT) primer [5', TGTTACCAATCGAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 23.9%; Score 405.8; DB 9; Length 428;
Best Local Similarity 98.1%; Pred. No. 1.8e-104;
Matches 421; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 353 CCCTCAGAGATATTTCCCAATTAATATACCTAAGTGAAGTCCCACTCTGCAACCCAA 412
Db CACTTGGAGATATATTTCCCAATTAATATACCTAAGTGAAGTCCCACTCTGCAACCCAA 369

QY 413 TGTCACTTCGAGTCTTACTGAACTACTTGAATTTCCGAGATCTCACTCTTCTTC 472
Db TGTCACTTCGAGTCTTACTGAACTACTTGAATTTCCGAGATCTCACTCTTCTTC 309

QY 473 GCCTGTACCTGTGCGGAAAGTCAAGCTTCCACTTCTCCCTGCTTCCACTCCCAAA 532
Db GCCTGTACCTGTGCGGAAAGTCAAGCTTCCACTTCTCCCTGCTTCCACTCCCAAA 249

QY 533 TACTTCGTGGTTTGCAGCTCTGGAGTATTTACCGTGTGGCTGTTAAATTTCTGCTC 592
Db TACTTCGTGGTTTGCAGCTCTGGAGTATTTACCGTGTGGCTGTTAAATTTCTGCTC 189

QY 593 CATCAGAAGCGAGAACTGACTCGGAACTATTTCATCCCGAGGATAGTAGAGCTTA 652
Db CATCAGAAGCGAGAACTGACTCGGAACTATTTCATCCCGAGGATAGTAGAGCTTA 129

QY 653 AAAAAAGCGAAGAGTGGGTGGGAGGACTTCAGTAAATCATCAGGTGGCAGCTCAAT 712
Db AAAAAAGCGAAGAGTGGGTGGGAGGACTTCAGTAAATCATCAGGTGGCAGCTCAAT 70

QY 713 TTATCTTTGTGAACGCTGGATAGTAATCCCTCTATCAGTGGCTGTTCAGGAATAAG 772
Db TTATCTTTGTGAACGCTGGATAGTAATCCCTCTATCAGTGGCTGTTCAGGAATAAG 10

QY 773 TGAATAAAC 781
Db 9 TGAATAAAC 1

RESULT 11
N67996/c
LOCUS
DEFINITION 412 bp mRNA linear EST 13-MAR-1996
yz92s04.s1 Soares_multiple_sclerosis_2NbHSP Homo sapiens cDNA

clone IMAGE:290526 3', mRNA sequence.

N67996
VERSION
N67996.1 GI:1224157
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 412)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL, contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: ml3 -40 forward
High quality sequence stop: 355.
Location/Qualifiers
1. 412
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3906366"
/db_xref="taxon:9606"
/clone="IMAGE:290526"
/sex="male"
/tissue_type="multiple sclerosis lesions"
/dev_stage="Age 46"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares multiple sclerosis 2NBHNSP"
/notes="Vector: pT73D (Pharmacia) with a modified
polylinker V type: phagemid; Site 1: Not I; Site 2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo (dT)
primer [5']
TGTTACCAATCTGAAGTGGAGCGCGCATTTTTTTTTTTTTTTT 3',
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M. Fatima Bonaudo. RNA from 4 multiple sclerosis
lesions from one patient was kindly provided by Dr. Kevin
G. Becker (NINDS/NIH)."

Query Match 23.5%; Score 399.4; DB 14; Length 412;
Best Local Similarity 99.5%; Pred. No. 1.2e-102;
Matches 411; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 411 AATGTCACATTCAGTCTTACTGAACTACTGACTGCAATTCCTCCAGATCTCACCTCTTC 470
DB 412 AATGTCACATTCAGTCTTACTGAACTACTGACTGCAATTCCTCCAGATCTCACCTCTTC 353
QY 471 TCGCCTGTACCTCTGCGGGAAGTCAGCCCTCCACCTTCCTCCCTGCTCCACTCCCAA 530
DB 352 TCGCCTGTACCTCTGCGGGAAGTCAGCCCTCCACCTTCCTCCCTGCTCCACTCCCAA 293
QY 531 AATACTTCGTGGTGTTCGACCTCTCGAGTATTTTACCGTGTGGCTGTAAATTTCTGCC 590
DB 292 AATACTTCGTGGTGTTCGACCTCTCGAGTATTTTACCGTGTGGCTGTAAATTTCTGCC 233
QY 591 TCCATCAGAAGGAGAACTGATTCGCGAATTTATTCATCCCGAGCCGATAGACGCT 650
DB 232 TCCATCAGAAGGAGAACTGATTCGCGAATTTATTCATCCCGAGCCGATAGACGCT 173

clone IMAGE:290526 3', mRNA sequence.

N67996
VERSION
N67996.1 GI:1224157
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 412)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL, contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: ml3 -40 forward
High quality sequence stop: 355.
Location/Qualifiers
1. 412
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3906366"
/db_xref="taxon:9606"
/clone="IMAGE:290526"
/sex="male"
/tissue_type="multiple sclerosis lesions"
/dev_stage="Age 46"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares multiple sclerosis 2NBHNSP"
/notes="Vector: pT73D (Pharmacia) with a modified
polylinker V type: phagemid; Site 1: Not I; Site 2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo (dT)
primer [5']
TGTTACCAATCTGAAGTGGAGCGCGCATTTTTTTTTTTTTTTT 3',
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M. Fatima Bonaudo. RNA from 4 multiple sclerosis
lesions from one patient was kindly provided by Dr. Kevin
G. Becker (NINDS/NIH)."

Query Match 23.5%; Score 399.4; DB 14; Length 412;
Best Local Similarity 99.5%; Pred. No. 1.2e-102;
Matches 411; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 411 AATGTCACATTCAGTCTTACTGAACTACTGACTGCAATTCCTCCAGATCTCACCTCTTC 470
DB 412 AATGTCACATTCAGTCTTACTGAACTACTGACTGCAATTCCTCCAGATCTCACCTCTTC 353
QY 471 TCGCCTGTACCTCTGCGGGAAGTCAGCCCTCCACCTTCCTCCCTGCTCCACTCCCAA 530
DB 352 TCGCCTGTACCTCTGCGGGAAGTCAGCCCTCCACCTTCCTCCCTGCTCCACTCCCAA 293
QY 531 AATACTTCGTGGTGTTCGACCTCTCGAGTATTTTACCGTGTGGCTGTAAATTTCTGCC 590
DB 292 AATACTTCGTGGTGTTCGACCTCTCGAGTATTTTACCGTGTGGCTGTAAATTTCTGCC 233
QY 591 TCCATCAGAAGGAGAACTGATTCGCGAATTTATTCATCCCGAGCCGATAGACGCT 650
DB 232 TCCATCAGAAGGAGAACTGATTCGCGAATTTATTCATCCCGAGCCGATAGACGCT 173

clone IMAGE:290526 3', mRNA sequence.

N67996
VERSION
N67996.1 GI:1224157
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 412)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL, contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: ml3 -40 forward
High quality sequence stop: 355.
Location/Qualifiers
1. 412
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3906366"
/db_xref="taxon:9606"
/clone="IMAGE:290526"
/sex="male"
/tissue_type="multiple sclerosis lesions"
/dev_stage="Age 46"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares multiple sclerosis 2NBHNSP"
/notes="Vector: pT73D (Pharmacia) with a modified
polylinker V type: phagemid; Site 1: Not I; Site 2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo (dT)
primer [5']
TGTTACCAATCTGAAGTGGAGCGCGCATTTTTTTTTTTTTTTT 3',
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M. Fatima Bonaudo. RNA from 4 multiple sclerosis
lesions from one patient was kindly provided by Dr. Kevin
G. Becker (NINDS/NIH)."

Query Match 23.5%; Score 399.4; DB 14; Length 412;
Best Local Similarity 99.5%; Pred. No. 1.2e-102;
Matches 411; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 411 AATGTCACATTCAGTCTTACTGAACTACTGACTGCAATTCCTCCAGATCTCACCTCTTC 470
DB 412 AATGTCACATTCAGTCTTACTGAACTACTGACTGCAATTCCTCCAGATCTCACCTCTTC 353
QY 471 TCGCCTGTACCTCTGCGGGAAGTCAGCCCTCCACCTTCCTCCCTGCTCCACTCCCAA 530
DB 352 TCGCCTGTACCTCTGCGGGAAGTCAGCCCTCCACCTTCCTCCCTGCTCCACTCCCAA 293
QY 531 AATACTTCGTGGTGTTCGACCTCTCGAGTATTTTACCGTGTGGCTGTAAATTTCTGCC 590
DB 292 AATACTTCGTGGTGTTCGACCTCTCGAGTATTTTACCGTGTGGCTGTAAATTTCTGCC 233
QY 591 TCCATCAGAAGGAGAACTGATTCGCGAATTTATTCATCCCGAGCCGATAGACGCT 650
DB 232 TCCATCAGAAGGAGAACTGATTCGCGAATTTATTCATCCCGAGCCGATAGACGCT 173

clone IMAGE:290526 3', mRNA sequence.

N67996
VERSION
N67996.1 GI:1224157
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 412)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL, contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: ml3 -40 forward
High quality sequence stop: 355.
Location/Qualifiers
1. 412
/organism="Homo sapiens"

the driver population), plus a pool of 11,136 clones from NCI CGAP Sub3 (IMAGE Clonetrds 2712456-2723591) (10% of the driver population), plus a pool of 5,472 clones from NCI CGAP Sub4 (IMAGE Clonetrds 2723592-2728959) (70% of the driver population). Subtraction was performed as previously described (Bonaldo, Lennon & Soares (1996)) : Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery. Genome Research 6, 791-806.

TAG TISSUE=germ cell
TAG LIB=NCI CGAP GC4
TAG SEQ=AAATC"

ORIGIN

Query Match	20.9%;	Score 354.4;	DB 10;	Length 371;
Best Local Similarity	99.7%;	Pred. No. 8.5e-90;		
Matches 355;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
858	TTGGAAGGAGACACATTTTTTAAATATATGCTGGTAAACCGGATCAGAAGGCAGGT	917		
Db	16	TTGGAAGGAGACACATTTTTTAAATATATGCTGGTAAACCGGATCAGAAGGCAGGT	75	
Qy	918	CCCCATGGAGCACACCCCTCGCCCTAAACATGCTGAACCCGGGCTGCCATAGCCTCGGTGG	977	
Db	76	CCCCATGGAGCACACCCCTCGCCCTAAACATGCTGAACCCGGGCTGCCATATCCTCGGTGG	135	
Qy	978	TCCTCCAAAGTGACTGCTCCGACAAAGGGTACGCTCTTCAAACGCATACGTTTAAAGC	1037	
Db	136	TCCTCCAAAGTGACTGCTCCGACAAAGGGTACGCTCTTCAAACGCATACGTTTAAAGC	195	
Qy	1038	AATTCCAGAAACCCCTCGGCTGTCCCGACTACACGGCCATTAAAGAAAGACGACTCTA	1097	
Db	196	AATTCCAGAAACCCCTCGGCTGTCCCGACTACACGGCCATTAAAGAAAGACGACTCTA	255	
Qy	1098	TGCCCGCCGTAATGTTTCTCAGATCACAGGACCCGTAATTTGGAGCTGGAGGGAGGGAAGC	1157	
Db	256	TGCCCGCCGTAATGTTTCTCAGATCACAGGACCCGTAATTTGGAGCTGGAGGGAGGGAAGC	315	
Qy	1158	CTTTTCTTCCAGGGGGCTAAGCGGCTTTCGAGCCCCCTTCCAAATCCCGGCTCCGG	1213	
Db	316	CTTTTCTTCCAGGGGGCTAAGCGGCTTTCGAGCCCCCTTCCAAATCCCGGCTCCGG	371	

RESULT 13
 BE858954/c
 LOCUS
 DBSFIT0N
 ACCESSION
 VERSION
 KEYWORDS
 ORGANISM
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 BE858954 378 bp mRNA linear EST 29-SEP-2000
 7944f12.x1 NCI_CGAP_P-28 Homo sapiens cDNA clone IMAGE:3309359 3',
 mRNA sequence.
 BE858954
 BE858954.1 GI:10374531
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 378)
 NCI_CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbsr@mail.nih.gov
 Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL, send email to:
info@image.llnl.gov
 Seq primer: -400P from Gibco
 High quality sequence stop: 342.
 Location/Qualifiers
 1. .378
 FEATURES
 source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3309359"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="NCI CGAP Pr28"
/note="Organ: prostate; Vector: pTTT3D-Pac (Pharmacia)
with a modified polylinker; Plasmid DNA from the
normalized library NCI CGAP Pr22 was prepared, and ss
circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (clonesIDs
985608-986759, 1101192-1101959, and 1217928-1220615).
Subtraction by Bento Soares and M. Fatima Bonaldo."

```

ORIGIN

	Query Match	20.9%;	Score 354.2;	DB 10;	Length 378;
	Best Local Similarity	97.6%;	Pred. No. 9.8e-90;		
	Matches 370;	Conservative 0;	Mismatches 8;	Indels 1;	Gaps 1;
QY	404	GCACCCCAATGTCACATTCGAGTCTTACTGAAGTACTCTTGCATGCGATTTCCCGAGATCTCA	463		
DB	378	GCACCCCAATGTCACATTCGAGTCTTACTGAAGTACTCTTGCATGCGATTTCCCGAGATCTCA	319		
QY	464	CTCTTCTCGCTGTACCCGTGTGCGCGAAAGTCAGCCCTCACCTTCTCCCTGCTTCCA	523		
DB	318	CTCTTCTCGCTGTACCCGTGTGCGCGAAAGTCAGCCCTCACCTTCTCCCTGCTTCCA	259		
QY	524	CTCCCAAAATATCTCGTGGTTTTCGAGCTCTGGAGTATTACCGTGTTGGCTGTTTAAAT	583		
DB	258	CTCCCAAAATATCTCGTGGTTTTCGAGCTATGGAGTATTTACCGTGTTGGCTGTTTAAAT	199		
QY	584	TTTCGCTCTCCATCAGAAGCGAAGAACTGACTCGCGAACTATTTCATCCCCAGCCGATAGT	643		
DB	198	TTTCGCTCTCCATCAGAAGCGAAGAACTGACTCGCGAACTATTTCATCCCCAGCCGATAGT	139		
QY	644	AGACGCTTAAAAAGAAACGGAAGAGGTGGTGGAGGACTTTCAGTAATCATCAGGTGGCA	703		
DB	138	AGACGCTTAAAAAGAAACGGAAGAGGTGGGT--GGAGGACTTTCAGTAATCATCAGGTGGCA	80		
QY	704	GCTCAATTTTATCGTTGTGTAACGTCGGATAGTAATCCCTCTATCATCGTGGCTGTTGCA	763		
DB	79	GCTCAATTTTATCGTTGTGTAACGTCGGATAGTAATCCCTCTATCATCGTGGCTGTTGCA	20		
QY	764	GGATAAAGTGAATAACA	782		
DB	19	GGATAAATAAAAAAGCCCA	1		

RESULT 14

BF570994 866 bp linear EST 12-DEC-2000
 602075915F1 NIH_MGC_62 Homo sapiens cDNA clone IMAGE:4243371 5',
 mRNA sequence.
 BF570994
 BF570994.1 GI:11644706
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 866)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 JOURNAL
 COMMENT

REFERENCE

AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC/DC/DTP cDNA Library Preparation: CLONETECH Laboratories, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

FEATURES

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

Db 241 TAGCGGAGCGGAGCGGACCCACAGAGCCCTGAGCAGCCCCACCGCCGCCGCCGCC 300
Qy 1667 TAGTTACCATCACACCCCGGAGGAGCGCGCAG 1698
Db 301 TAGTTACCATCACACCCCGGAGGAGCGCGCAG 332

Search completed: March 28, 2004, 17:06:24
Job time : 4645 secs